JOURNAL OF Economic Theory

How To Comply With The New Copyright Law

Libraries everywhere have found the easy way to fill photocopy requests legally and instantly, without the need to seek permissions, from this and over 3000 other key publications in business, science, humanities, and social science.

Participation in the Copyright Clearance Center (CCC) assures you of legal photocopying at the moment of need. You can:

Fill requests for multiple copies, interlibrary loan (beyond the CONTU guidelines), and reserve desk without fear of copyright infringement.

Supply copies simply and easily from registered publications. The CCC's flexible reporting system accepts photocopying reports and returns an itemized invoice. *You need not keep any records*, our computer will do it for you.

The Copyright Clearance Center is your one-stop place for on-the-spot clearance to photocopy for internal use. You will never have to decline a photocopy request or wonder about compliance with the law for any publication registered with the CCC.

For more information, just contact:

Copyright Clearance Center
21 Congress Street
Salem, Massachusetts 01970
(617) 744-3350

a not-for-profit corporation

NAME TITLE

ORGANIZATION

ADDRESS

CITY STATE ZIP

COUNTRY TELEPHONE

Journal of Economic Theory

Editor: Karl Shell
CARESS, Department of Economics
University of Pennsylvania
3718 Locust Walk/CR
Philadelphia PA 19104 USA

Associate Editors:

Yves Balasko
CEPREMAP

William A. Brock
University of Chicago and
University of Wisconsin

David Cass
University of Pennsylvania

Paul Champsaur
INSEE

Gerard Debreu
University of California at
Berkeley

Peter C. Fishburn
Bell Laboratories

Richard J. Gilbert
University of California at
Berkeley

Steven M. Goldman
University of California at
Berkeley

Jean-Michel Grandmont
CEPREMAP

Sanford Grossman
University of Pennsylvania

Roger Guesnerie
CEPREMAP and EHESS

Walter Perrin Heller
University of California at
San Diego

Werner Hildenbrand
Universität Bonn

Ehud Kalai
Northwestern University

David M. Kreps
Stanford University

Mordecai Kurz
Stanford University

Jean-Jacques Laffont
Université Toulouse I

Mukul Majumdar
Cornell University

Edmond Malinvaud
INSEE

Prasanta K. Pattanaik
University of Birmingham

Roy Radner
Bell Laboratories

John Roberts
Northwestern University

Stephen A. Ross
Yale University

Michael Rothschild
University of Wisconsin

Wayne J. Shafer
University of Southern
California

David A. Starrett
Stanford University

Charles A. Wilson
University of Wisconsin

Assistant Editor:

Susan Witherow Schulze
Journal of Economic Theory
University of Pennsylvania
3718 Locust Walk/CR
Philadelphia PA 19104 USA
Telephone: (215) 243-8487

Journal of Economic Theory
Volume 24, Number 1, February 1981

Contents

MENAHEM E. YAARI. Rawls, Edgeworth, Shapley, Nash: Theories of Distributive Justice Re-examined. 1

KI HANG KIM AND FRED W. ROUSH. Effective Nondictatorial Domains. 40

FUMITAKA SATO. On the Informational Size of Message Spaces for Resource Allocation Processes in Economies with Public Goods. 48

DAVID P. BARON AND RAYMOND R. DE BONDT. On the Design of Regulatory Price Adjustment Mechanisms. 70

CHARLES A. WILSON. Equilibrium in Dynamic Models with an Infinity of Agents. 95

YVES BALASKO AND KARL SHELL. The Overlapping-Generations Model. II. The Case of Pure Exchange with Money. 112

YVES BALASKO AND KARL SHELL. The Overlapping-Generations Model. III. The Case of Log-Linear Utility Functions. 143

ACKNOWLEDGMENT. 153

PAPERS TO APPEAR IN FORTHCOMING ISSUES. 155

Information for Authors

The *Journal of Economic Theory* will publish original articles on economic theory and related mathematical techniques. Manuscripts should be submitted in triplicate, to *Journal of Economic Theory*, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia PA 19104 USA.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be published nor is presently submitted elsewhere. While under editorial review, *it is the responsibility of the authors to keep the Editor informed about submissions, publication plans, and publication of related research (or abstracts thereof) in other outlets—including letters journals, journals in other disciplines, collections of articles, and published dissertations.*

It is understood that all persons listed as authors have given their approval for the submission of the paper; further, that any person cited as a source of personal communication has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other materials published in the *Journal of Economic Theory* represent the opinions of the author(s) and should not be construed to reflect the opinion of the Editorial Board or the Publisher.

Manuscripts should be prepared according to the following style rules (deviation from these rules causes publication delays):

Form of Manuscript. Submit manuscripts in triplicate and double-spaced on one side of 8.5 × 11 inch white paper. Number each page. *Page 1* should contain the article title, author and coauthor names, and complete affiliation(s) (name of institution, city, state, and zip code). At the bottom of page 1 place any footnotes to the title (indicated by superscript *, †, ‡). *Page 2* should contain a proposed running head (abbreviated form of the title) of less than 35 characters. It should also contain the name and mailing address of the author to whom proofs should be sent. *Page 3* should contain an abstract of no more than 100 words. The Abstract will appear at the beginning of the article in the journal and will be published in the *Journal of Economic Literature*. Use the abstract format which is required by the *Journal of Economic Literature*.

List of Symbols. Attach to the manuscript a complete typewritten list of symbols, identified typographically, not mathematically. This list will not appear in print but is essential in order to avoid costly author's corrections in proof. [If equations are handwritten in the text then the list of symbols should also be handwritten.] Distinguish between "oh," "zero"; "el," "one"; "kappa," "kay"; upper and lower case "kay"; etc. Indicate also when special type is required (German, Greek, vector, scalar, script, etc.); all other letters will be set in *italic*.

Tables. Number tables consecutively with Roman numerals. Extensive tables will be reproduced as photoengravings and so should be typed carefully and in the *exact* format desired. Authors will be charged for any new engravings necessitated by changes in proof. Use superscript lower-case italic letters (*a, b, c*) for table footnotes, which should be typed immediately below the table. Type tables at least double-spaced, including titles and footnotes. Do *not* underline table titles; reserve underlining for text that is to be italicized.

Equations should be typewritten and with the number placed in parentheses at the right margin. Reference to equations should use the form "Eq. (3)" or simply (3). Superscripts and subscripts should be typed or handwritten clearly above and below the line, respectively. Use the exponent $^{1/2}$ wherever possible.

"Style." In general, authors should be guided by *A Manual for Authors* ($1.00) published in 1962 (and revised in 1966) by the American Mathematical Society, P.O. Box 6248, Providence, R. I. 02904, **but**

References should be styled and punctuated according to the following examples:

1. G. G. Lorentz, The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239–251. [Do not underline titles of articles or books.]
2. J. T. Scheick, Some problems in approximation theory, Ph. D. Dissertation, Syracuse University, June, 1966.
3. E. Hille and R. Phillips, "Functional Analysis and Semi-groups," American Mathematical Society, Providence, R. I., 1957. [For books, include name of publisher, city, state, and year.]

For unpublished lectures or symposia, include title of the paper, name of the sponsoring society in full, and date. For journal names, follow *Abbreviations of the Names of Scientific Periodicals Reviewed in Mathematical Reviews*, Volume 37, 1969 (reprint $1.00, American Mathematical Society). Cite references in the text by an Arabic number between square brackets, as [1], [1, 2], [1, Theorem 1.5], etc. Type references double--spaced throughout.

Footnotes in text should be avoided if at all possible. If they must be used, identify by superscript numbers and type together on a separate page, double-spaced.

Figures. All illustrations are to be considered as figures. Number each graph or drawing in sequence with Arabic numerals. Supply a descriptive legend for each figure. Type legends double- or triple-spaced consecutively on a separate sheet. **Figures must be submitted in a form suitable for reproduction.** Plan figures to fit the proportion of the printed page. Use a professional lettering set on the original so that the letters and numbers are large enough and "open" enough to take a reduction of 50–60% without filling in with ink. Do not include background grids; however, on paper with blue lines the grid can be eliminated in the process of photoreproduction. Identify each figure in a margin with the name of the journal, author's name, and figure number; avoid marking the backs of figures.

Proofs. Galley proofs will be sent to the author, with a reprint order form. Authors will be charged for alterations in excess of 10% of the cost of composition.

Reprints. Fifty reprints without covers will be provided free of charge. Additional reprints may be purchased; an order form will be included with proofs.

Journal of Economic Theory
Volume 24, Number 2, April 1981

Contents

KIYOHIKO GIICHI NISHIMURA. On Uniqueness of a Steady State and Convergence of Optimal Paths in Multisector Models of Optimal Growth with a Discount Rate. 157

L. C. MACLEAN, C. A. FIELD, AND W. R. S. SUTHERLAND. Optimal Growth and Uncertainty: The Borrowing Models. 168

KOFI O. NTI AND MARTIN SHUBIK. Noncooperative Oligopoly with Entry. 187

FRANK MILNE. Induced Preferences and the Theory of the Consumer. 205

M. ALI KHAN AND AKIRA YAMAZAKI. On the Cores of Economies with Indivisible Commodities and a Continuum of Traders. . . 218

PETER COUGHLIN AND SHMUEL NITZAN. Directional and Local Electoral Equilibria with Probabilistic Voting. 226

GEOFFREY LEWIS. The Phillips Curve and Bayesian Learning. . . . 240

TRUMAN F. BEWLEY. Stationary Equilibrium. 265

NOTES, COMMENTS, AND LETTERS TO THE EDITOR.

XAVIER FREIXAS. Optimal Growth with Experimentation. . . . 296

PAPERS TO APPEAR IN FORTHCOMING ISSUES. 310

Information for Authors

The *Journal of Economic Theory* will publish original articles on economic theory and related mathematical techniques. Manuscripts should be submitted in triplicate, to *Journal of Economic Theory,* University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia PA 19104 USA.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be published nor is presently submitted elsewhere. While under editorial review, *it is the responsibility of the authors to keep the Editor informed about submissions, publication plans, and publication of related research (or abstracts thereof) in other outlets—including letters journals, journals in other disciplines, collections of articles, and published dissertations.*

It is understood that all persons listed as authors have given their approval for the submission of the paper; further, that any person cited as a source of personal communication has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other materials published in the *Journal of Economic Theory* represent the opinions of the author(s) and should not be construed to reflect the opinion of the Editorial Board or the Publisher.

Manuscripts should be prepared according to the following style rules (deviation from these rules causes publication delays):

Form of Manuscript. Submit manuscripts in triplicate and double-spaced on one side of 8.5 × 11 inch white paper. Number each page. *Page 1* should contain the article title, author and coauthor names, and complete affiliation(s) (name of institution, city, state, and zip code). At the bottom of page 1 place any footnotes to the title (indicated by superscript *, †, ‡). *Page 2* should contain a proposed running head (abbreviated form of the title) of less than 35 characters. It should also contain the name and mailing address of the author to whom proofs should be sent. *Page 3* should contain an abstract of no more than 100 words. The Abstract will appear at the beginning of the article in the journal and will be published in the *Journal of Economic Literature.* Use the abstract format which is required by the *Journal of Economic Literature.*

List of Symbols. Attach to the manuscript a complete typewritten list of symbols, identified typographically, not mathematically. This list will not appear in print but is essential in order to avoid costly author's corrections in proof. [If equations are handwritten in the text then the list of symbols should also be handwritten.] Distinguish between "oh," "zero"; "el," "one"; "kappa," "kay"; upper and lower case "kay"; etc. Indicate also when special type is required (German, Greek, vector, scalar, script, etc.); all other letters will be set in *italic.*

Tables. Number tables consecutively with Roman numerals. Extensive tables will be reproduced as photoengravings and so should be typed carefully and in the *exact* format desired. Authors will be charged for any new engravings necessitated by changes in proof. Use superscript lower-case italic letters (*a*, *b*, *c*) for table footnotes, which should be typed immediately below the table. Type tables at least double-spaced, including titles and footnotes. Do *not* underline table titles; reserve underlining for text that is to be italicized.

Equations should be typewritten and with the number placed in parentheses at the right margin. Reference to equations should use the form "Eq. (3)" or simply (3). Superscripts and subscripts should be typed or handwritten clearly above and below the line, respectively. Use the exponent $^{1/2}$ wherever possible.

"Style." In general, authors should be guided by *A Manual for Authors* ($1.00) published in 1962 (and revised in 1966) by the American Mathematical Society, P.O. Box 6248, Providence, R. I. 02904, **but**

References should be styled and punctuated according to the following examples:

1. G. G. Lorentz, The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239–251. [Do not underline titles of articles or books.]
2. J. T. Scheick, Some problems in approximation theory, Ph. D. Dissertation, Syracuse University, June, 1966.
3. E. Hille and R. Phillips, "Functional Analysis and Semi-groups," American Mathematical Society, Providence, R. I., 1957. [For books, include name of publisher, city, state, and year.]

For unpublished lectures or symposia, include title of the paper, name of the sponsoring society in full, and date. For journal names, follow *Abbreviations of the Names of Scientific Periodicals Reviewed in Mathematical Reviews,* Volume 37, 1969 (reprint $1.00, American Mathematical Society). Cite references in the text by an Arabic number between square brackets, as [1], [1, 2], [1, Theorem 1.5], etc. Type references double--spaced throughout.

Footnotes in text should be avoided if at all possible. If they must be used, identify by superscript numbers and type together on a separate page, double-spaced.

Figures. All illustrations are to be considered as figures. Number each graph or drawing in sequence with Arabic numerals. Supply a descriptive legend for each figure. Type legends double- or triple-spaced consecutively on a separate sheet. **Figures must be submitted in a form suitable for reproduction.** Plan figures to fit the proportion of the printed page. Use a professional lettering set on the original so that the letters and numbers are large enough and "open" enough to take a reduction of 50–60% without filling in with ink. Do not include background grids; however, on paper with blue lines the grid can be eliminated in the process of photoreproduction. Identify each figure in a margin with the name of the journal, author's name, and figure number; avoid marking the backs of figures.

Proofs. Galley proofs will be sent to the author, with a reprint order form. Authors will be charged for alterations in excess of 10% of the cost of composition.

Reprints. Fifty reprints without covers will be provided free of charge. Additional reprints may be purchased; an order form will be included with proofs.

Journal of Economic Theory
Volume 24, Number 3, June 1981

Contents

MAMORU KANEKO. A Bilateral Monopoly and the Nash Cooperative Solution. 311

CHIN LIM. Theory of the Firm: Uncertainty and Choice of Experiments. 328

KENNETH BURDETT AND DAVID A. MALUEG. The Theory of Search for Several Goods. 362

JOHN LAITNER. The Steady States of a Stochastic Decentralized Growth Model. 377

HANS KEIDING. Existence of Budget Constrained Pareto Efficient Allocations. 393

HERVÉ MOULIN. Prudence versus Sophistication in Voting Strategy. 398

CLAUDE FOURGEAUD, BERNARD LENCLUD, AND PHILIPPE MICHEL. Two-Sector Model with Quantity Rationing. 413

NOTES, COMMENTS, AND LETTERS TO THE EDITOR.

ARTHUR J. ROBSON. Sufficiency of the Pontryagin Conditions for Optimal Control when the Time Horizon Is Free. . . . 437

MORGAN ÅBERG AND HÅKAN PERSSON. A Note on a Closed Input–Output Model with Finite Life-Times and Gestation Lags. 446

D. J. HARTFIEL. Bounds on the Solutions to $x' = SAx$ where A is S-Stable. 453

ANNOUNCEMENT. 458

PAPERS TO APPEAR IN FORTHCOMING ISSUES. 459

AUTHOR INDEX FOR VOLUME 24. 460

Information for Authors

The *Journal of Economic Theory* will publish original articles on economic theory and related mathematical techniques. Manuscripts should be submitted in triplicate, to *Journal of Economic Theory*, University of Pennsylvania, 3718 Locust Walk CR, Philadelphia PA 19104 USA.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be published nor is presently submitted elsewhere. While under editorial review, *it is the responsibility of the authors to keep the Editor informed about submissions, publication plans, and publication of related research (or abstracts thereof) in other outlets—including letters journals, journals in other disciplines, collections of articles, and published dissertations.*

It is understood that all persons listed as authors have given their approval for the submission of the paper; further, that any person cited as a source of personal communication has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other materials published in the *Journal of Economic Theory* represent the opinions of the author(s) and should not be construed to reflect the opinion of the Editorial Board or the Publisher.

Manuscripts should be prepared according to the following style rules (deviation from these rules causes publication delays):

Form of Manuscript. Submit manuscripts in triplicate and double-spaced on one side of 8.5 × 11 inch white paper. Number each page. *Page 1* should contain the article title, author and coauthor names, and complete affiliation(s) (name of institution, city, state, and zip code). At the bottom of page 1 place any footnotes to the title (indicated by superscript *, †, ‡). *Page 2* should contain a proposed running head (abbreviated form of the title) of less than 35 characters. It should also contain the name and mailing address of the author to whom proofs should be sent. *Page 3* should contain an abstract of no more than 100 words. The Abstract will appear at the beginning of the article in the journal and will be published in the *Journal of Economic Literature.* Use the abstract format which is required by the *Journal of Economic Literature.*

List of Symbols. Attach to the manuscript a complete typewritten list of symbols, identified typographically, not mathematically. This list will not appear in print but is essential in order to avoid costly author's corrections in proof. [If equations are handwritten in the text then the list of symbols should also be handwritten.] Distinguish between "oh," "zero"; "el," "one"; "kappa," "kay"; upper and lower case "kay"; etc. Indicate also when special type is required (German, Greek, vector, scalar, script, etc.); all other letters will be set in *italic.*

Tables. Number tables consecutively with Roman numerals. Extensive tables will be reproduced as photoengravings and so should be typed carefully and in the *exact* format desired. Authors will be charged for any new engravings necessitated by changes in proof. Use superscript lower-case italic letters (*a*, *b*, *c*) for table footnotes, which should be typed immediately below the table. Type tables at least double-spaced, including titles and footnotes. Do *not* underline table titles; reserve underlining for text that is to be italicized.

Equations should be typewritten and with the number placed in parentheses at the right margin. Reference to equations should use the form "Eq. (3)" or simply (3). Superscripts and subscripts should be typed or handwritten clearly above and below the line, respectively. Use the exponent $^{1/2}$ wherever possible.

"Style." In general, authors should be guided by *A Manual for Authors* ($1.00) published in 1962 (and revised in 1966) by the American Mathematical Society, P.O. Box 6248, Providence, R. I. 02904, **but**

References should be styled and punctuated according to the following examples:

1. G. G. Lorentz, The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239–251. [Do not underline titles of articles or books.]
2. J. T. Scheick, Some problems in approximation theory, Ph. D. Dissertation, Syracuse University, June, 1966.
3. E. Hille and R. Phillips, "Functional Analysis and Semi-groups," American Mathematical Society, Providence, R. I., 1957. [For books, include name of publisher, city, state, and year.]

For unpublished lectures or symposia, include title of the paper, name of the sponsoring society in full, and date. For journal names, follow *Abbreviations of the Names of Scientific Periodicals Reviewed in Mathematical Reviews*, Volume 37, 1969 (reprint $1.00, American Mathematical Society). Cite references in the text by an Arabic number between square brackets, as [1], [1, 2], [1, Theorem 1.5], etc. Type references double- or triple-spaced throughout.

Footnotes in text should be avoided if at all possible. If they must be used, identify by superscript numbers and type together on a separate page, double- or triple-spaced.

Figures. All illustrations are to be considered as figures. Number each graph or drawing in sequence with Arabic numerals. Supply a descriptive legend for each figure. Type legends double- or triple-spaced consecutively on a separate sheet. **Figures must be submitted in a form suitable for reproduction.** Plan figures to fit the proportion of the printed page. Use a professional lettering set on the original so that the letters and numbers are large enough and "open" enough to take a reduction of 50–60% without filling in with ink. Do not include background grids; however, on paper with blue lines the grid can be eliminated in the process of photoreproduction. Identify each figure in ı margin with the name of the journal, author's name, and figure number; avoid marking the backs of igures.

Proofs. Galley proofs will be sent to the author, with a reprint order form. Authors will be charged or alterations in excess of 10% of the cost of composition.

Reprints. Fifty reprints without covers will be provided free of charge. Additional reprints may be urchased; an order form will be included with proofs.

Journal of Economic Theory

Volume 25, Number 1, August 1981

Contents

DONALD B. ROSENFIELD AND ROY D. SHAPIRO. Optimal Adaptive Price Search. 1

JENNIFER F. REINGANUM. Dynamic Games of Innovation. 21

K. BURDETT AND DALE T. MORTENSEN. Testing for Ability in a Competitive Labor Market. 42

ROGER GUESNERIE AND CLAUDE ODDOU. Second-Best Taxation as a Game. 67

ROBERT W. ROSENTHAL. Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox. 92

RONALD A. HEINER. Length and Cycle Equalization. 101

NOTES, COMMENTS, AND LETTERS TO THE EDITOR.

- FRANK A. COWELL AND KIYOSHI KUGA. Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement. 131
- MYRNA WOODERS. Equilibria, the Core and Jurisdiction Structures in Economies with a Local Public Good: A Correction. 144
- GARY CHAMBERLAIN AND MICHAEL ROTHSCHILD. A Note on the Probability of Casting a Decisive Vote. 152

PAPERS TO APPEAR IN FORTHCOMING ISSUES. 163

Information for Authors

The *Journal of Economic Theory* will publish original articles on economic theory and related mathematical techniques. Manuscripts should be submitted in triplicate, to *Journal of Economic Theory*, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia PA 19104 USA.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be published nor is presently submitted elsewhere. While under editorial review, *it is the responsibility of the authors to keep the Editor informed about submissions, publication plans, and publication of related research (or abstracts thereof) in other outlets—including letters journals, journals in other disciplines, collections of articles, and published dissertations.*

It is understood that all persons listed as authors have given their approval for the submission of the paper; further, that any person cited as a source of personal communication has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other materials published in the *Journal of Economic Theory* represent the opinions of the author(s) and should not be construed to reflect the opinion of the Editorial Board or the Publisher.

Manuscripts should be prepared according to the following style rules (deviation from these rules causes publication delays):

Form of Manuscript. Submit manuscripts in triplicate and double-spaced on one side of 8.5 × 11 inch white paper. Number each page. *Page 1* should contain the article title, author and coauthor names, and complete affiliation(s) (name of institution, city, state, and zip code). At the bottom of page 1 place any footnotes to the title (indicated by superscript *, †, ‡). *Page 2* should contain a proposed running head (abbreviated form of the title) of less than 35 characters. It should also contain the name and mailing address of the author to whom proofs should be sent. *Page 3* should contain an abstract of no more than 100 words. The Abstract will appear at the beginning of the article in the journal and will be published in the *Journal of Economic Literature*. Use the abstract format which is required by the *Journal of Economic Literature*.

List of Symbols. Attach to the manuscript a complete typewritten list of symbols, identified typographically, not mathematically. This list will not appear in print but is essential in order to avoid costly author's corrections in proof. [If equations are handwritten in the text then the list of symbols should also be handwritten.] Distinguish between "oh," "zero"; "el," "one"; "kappa," "kay"; upper and lower case "kay"; etc. Indicate also when special type is required (German, Greek, vector, scalar, script, etc.); all other letters will be set in *italic*.

Tables. Number tables consecutively with Roman numerals. Extensive tables will be reproduced as photoengravings and so should be typed carefully and in the *exact* format desired. Authors will be charged for any new engravings necessitated by changes in proof. Use superscript lower-case italic letters (*a*, *b*, *c*) for table footnotes, which should be typed immediately below the table. Type tables at least double-spaced, including titles and footnotes. Do *not* underline table titles; reserve underlining for text that is to be italicized.

Equations should be typewritten and with the number placed in parentheses at the right margin. Reference to equations should use the form "Eq. (3)" or simply (3). Superscripts and subscripts should be typed or handwritten clearly above and below the line, respectively. Use the exponent $^{1/2}$ wherever possible.

"Style." In general, authors should be guided by *A Manual for Authors* ($1.00) published in 1962 (and revised in 1966) by the American Mathematical Society, P.O. Box 6248, Providence, R. I. 02904, **but**

References should be styled and punctuated according to the following examples:

1. G. G. Lorentz, The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239–251. [Do not underline titles of articles or books.]
2. J. T. Scheick, Some problems in approximation theory, Ph. D. Dissertation, Syracuse University, June, 1966.
3. E. Hille and R. Phillips, "Functional Analysis and Semi-groups," American Mathematical Society, Providence, R. I., 1957. [For books, include name of publisher, city, state, and year.]

For unpublished lectures or symposia, include title of the paper, name of the sponsoring society in full, and date. For journal names, follow *Abbreviations of the Names of Scientific Periodicals Reviewed in Mathematical Reviews*, Volume 37, 1969 (reprint $1.00, American Mathematical Society). Cite references in the text by an Arabic number between square brackets, as [1], [1, 2], [1, Theorem 1.5], etc. Type references double-spaced throughout.

Footnotes in text should be avoided if at all possible. If they must be used, identify by superscript numbers and type together on a separate page, double-spaced.

Figures. All illustrations are to be considered as figures. Number each graph or drawing in sequence with Arabic numerals. Supply a descriptive legend for each figure. Type legends double- or triple-spaced consecutively on a separate sheet. **Figures must be submitted in a form suitable for reproduction.** Plan figures to fit the proportion of the printed page. Use a professional lettering set on the original so that the letters and numbers are large enough and "open" enough to take a reduction of 50–60% without filling in with ink. Do not include background grids; however, on paper with blue lines the grid can be eliminated in the process of photoreproduction. Identify each figure in a margin with the name of the journal, author's name, and figure number; avoid marking the backs of igures.

Proofs. Galley proofs will be sent to the author, with a reprint order form. Authors will be charged or alterations in excess of 10% of the cost of composition.

Reprints. Fifty reprints without covers will be provided free of charge. Additional reprints may be urchased; an order form will be included with proofs.

Journal of Economic Theory
Volume 25, Number 2, October 1981

Contents

P. A. DIAMOND AND ERIC MASKIN. An Equilibrium Analysis of Search and Breach of Contract. II. A Non-Steady State Example 165

YVES BALCER. Equilibrium Distributions of Sales and Advertising Prices over Space 196

TROUT RADER. Utility over Time: The Homothetic Case 219

W. HAQUE. Direct and Indirect Weak Separability 237

PETER C. FISHBURN. Majority Committees 255

PAUL CHAMPSAUR AND GUY LAROQUE. Fair Allocations in Large Economies 269

NOTES, COMMENTS, AND LETTERS TO THE EDITOR

TATSURO ICHIISHI. Super-modularity: Applications to Convex Games and to the Greedy Algorithm for *LP* 283

JOSEPH P. HUGHES. Giffen Inputs and the Theory of Multiple Production 287

JOHN C. ECKALBAR. Stable Quantities in Fixed Price Disequilibrium 302

ROSS M. STARR. Approximation of Points of the Convex Hull of a Sum of Sets by Points of the Sum: An Elementary Approach 314

ELIAKIM KATZ. A Note on a Comparative Statics Theorem for Choice under Risk 318

PAPERS TO APPEAR IN FORTHCOMING ISSUES 320

Information for Authors

The *Journal of Economic Theory* will publish original articles on economic theory and related mathematical techniques. Manuscripts should be submitted in triplicate, to *Journal of Economic Theory*, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia PA 19104 USA.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be published nor is presently submitted elsewhere. While under editorial review, *it is the responsibility of the authors to keep the Editor informed about submissions, publication plans, and publication of related research (or abstracts thereof) in other outlets—including letters journals, journals in other disciplines, collections of articles, and published dissertations.*

It is understood that all persons listed as authors have given their approval for the submission of the paper; further, that any person cited as a source of personal communication has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other materials published in the *Journal of Economic Theory* represent the opinions of the author(s) and should not be construed to reflect the opinion of the Editorial Board or the Publisher.

Manuscripts should be prepared according to the following style rules (deviation from these rules causes publication delays):

Form of Manuscript. Submit manuscripts in triplicate and double-spaced on one side of 8.5 × 11 inch white paper. Number each page. *Page 1* should contain the article title, author and coauthor names, and complete affiliation(s) (name of institution, city, state, and zip code). At the bottom of page 1 place any footnotes to the title (indicated by superscript *, †, ‡). *Page 2* should contain a proposed running head (abbreviated form of the title) of less than 35 characters. It should also contain the name and mailing address of the author to whom proofs should be sent. *Page 3* should contain an abstract of no more than 100 words. The Abstract will appear at the beginning of the article in the journal and will be published in the *Journal of Economic Literature*. Use the abstract format which is required by the *Journal of Economic Literature*.

List of Symbols. Attach to the manuscript a complete typewritten list of symbols, identified typographically, not mathematically. This list will not appear in print but is essential in order to avoid costly author's corrections in proof. [If equations are handwritten in the text then the list of symbols should also be handwritten.] Distinguish between "oh," "zero"; "el," "one"; "kappa," "kay"; upper and lower case "kay"; etc. Indicate also when special type is required (German, Greek, vector, scalar, script, etc.); all other letters will be set in *italic*.

Tables. Number tables consecutively with Roman numerals. Extensive tables will be reproduced as photoengravings and so should be typed carefully and in the *exact* format desired. Authors will be charged for any new engravings necessitated by changes in proof. Use superscript lower-case italic letters (a, b, c) for table footnotes, which should be typed immediately below the table. Type tables at least double-spaced, including titles and footnotes. Do *not* underline table titles; reserve underlining for text that is to be italicized.

Equations should be typewritten and with the number placed in parentheses at the right margin. Reference to equations should use the form "Eq. (3)" or simply (3). Superscripts and subscripts should be typed or handwritten clearly above and below the line, respectively. Use the exponent $^{1/2}$ wherever possible.

"Style." In general, authors should be guided by *A Manual for Authors* ($1.00) published in 1962 (and revised in 1966) by the American Mathematical Society, P.O. Box 6248, Providence, R. I. 02904, **but**

References should be styled and punctuated according to the following examples:

1. G. G. Lorentz, The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239–251. [Do not underline titles of articles or books.]
2. J. T. Scheick, Some problems in approximation theory, Ph. D. Dissertation, Syracuse University, June, 1966.
3. E. Hille and R. Phillips, "Functional Analysis and Semi-groups," American Mathematical Society, Providence, R. I., 1957. [For books, include name of publisher, city, state, and year.]

For unpublished lectures or symposia, include title of the paper, name of the sponsoring society in full, and date. For journal names, follow *Abbreviations of the Names of Scientific Periodicals Reviewed in Mathematical Reviews*, Volume 37, 1969 (reprint $1.00, American Mathematical Society). Cite references in the text by an Arabic number between square brackets, as [1], [1, 2], [1, Theorem 1.5], etc. Type references double-spaced throughout.

Footnotes in text should be avoided if at all possible. If they must be used, identify by superscript numbers and type together on a separate page, double-spaced.

Figures. All illustrations are to be considered as figures. Number each graph or drawing in sequence with Arabic numerals. Supply a descriptive legend for each figure. Type legends double- or triple-spaced consecutively on a separate sheet. **Figures must be submitted in a form suitable for reproduction.** Plan figures to fit the proportion of the printed page. Use a professional lettering set on the original so that the letters and numbers are large enough and "open" enough to take a reduction of 50–60% without filling in with ink. Do not include background grids; however, on paper with blue lines the grid can be eliminated in the process of photoreproduction. Identify each figure in a margin with the name of the journal, author's name, and figure number; avoid marking the backs of figures.

Proofs. Galley proofs will be sent to the author, with a reprint order form. Authors will be charged for alterations in excess of 10% of the cost of composition.

Reprints. Fifty reprints without covers will be provided free of charge. Additional reprints may be purchased; an order form will be included with proofs.

Journal of Economic Theory
Volume 25, Number 3, December 198

Contents

HERAKLIS M. POLEMARCHAKIS. Constrained Excess Demand Functions 323

I. MACINTYRE AND PRASANTA K. PATTANAIK. Strategic Voting under Minimally Binary Group Decision Functions 338

ROY GARDNER. Wealth and Power in a Collegial Polity 353

KOTARO SUZUMURA. On Pareto-Efficiency and the No-Envy Concept of Equity 367

COSTAS AZARIADIS. Self-Fulfilling Prophecies 380

W. E. DIEWERT, M. AVRIEL, AND I. ZANG. Nine Kinds of Quasiconcavity and Concavity 397

NEIL WALLACE. A Hybrid Fiat—Commodity Monetary System . . 421

NOTES, COMMENTS, AND LETTERS TO THE EDITOR.

WILLIAM THOMSON. A Class of Solutions to Bargaining Problems 431

R. MANNING. A Nonsubstitution Theorem with Many Primary Factors 442

RAYMOND J. KOPP. Measuring the Technical Efficiency of Production: A Comment 450

ROLF FÄRE AND C. A. KNOX LOVELL. Measuring the Technical Efficiency of Production: Reply 453

ROLF FÄRE AND DANIEL PRIMONT. Separability vs Strict Separability: A Further Result 455

YUK-SHEE CHAN. A Note on Risk and the Value of Information 461

GEORGE A. KONDOR. On the Asymptotic Local Stability of Autonomous Economic Systems with Nonnegativity Conditions 466

ERRATUM

Volume 24, Number 1, 1981: Yves Balasko and Karl Shell, "The Overlapping Generations Model. II. The Case of Pure Exchange with Money," pp. 112–142 471

PAPERS TO APPEAR IN FORTHCOMING ISSUES 472

AUTHOR INDEX FOR VOLUME 25 473

Information for Authors

The *Journal of Economic Theory* will publish original articles on economic theory and related mathematical techniques. Manuscripts should be submitted in triplicate, to *Journal of Economic Theory*, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia PA 19104 USA.

Manuscripts are accepted for review with the understanding that the same work has not been and will not be published nor is presently submitted elsewhere. While under editorial review, *it is the responsibility of the authors to keep the Editor informed about submissions, publication plans, and publication of related research (or abstracts thereof) in other outlets—including letters journals, journals in other disciplines, collections of articles, and published dissertations.*

It is understood that all persons listed as authors have given their approval for the submission of the paper; further, that any person cited as a source of personal communication has approved such citation. Written authorization may be required at the Editor's discretion. Articles and any other materials published in the *Journal of Economic Theory* represent the opinions of the author(s) and should not be construed to reflect the opinion of the Editorial Board or the Publisher.

Manuscripts should be prepared according to the following style rules (deviation from these rules causes publication delays):

Form of Manuscript. Submit manuscripts in triplicate and double-spaced on one side of 8.5 × 11 inch white paper. Number each page. *Page 1* should contain the article title, author and coauthor names, and complete affiliation(s) (name of institution, city, state, and zip code). At the bottom of page 1 place any footnotes to the title (indicated by superscript *, †, ‡). *Page 2* should contain a proposed running head (abbreviated form of the title) of less than 35 characters. It should also contain the name and mailing address of the author to whom proofs should be sent. *Page 3* should contain an abstract of no more than 100 words. The Abstract will appear at the beginning of the article in the journal and will be published in the *Journal of Economic Literature*. Use the abstract format which is required by the *Journal of Economic Literature*.

List of Symbols. Attach to the manuscript a complete typewritten list of symbols, identified typographically, not mathematically. This list will not appear in print but is essential in order to avoid costly author's corrections in proof. [If equations are handwritten in the text then the list of symbols should also be handwritten.] Distinguish between "oh," "zero"; "el," "one"; "kappa," "kay"; upper and lower case "kay"; etc. Indicate also when special type is required (German, Greek, vector, scalar, script, etc.); all other letters will be set in *italic*.

Tables. Number tables consecutively with Roman numerals. Extensive tables will be reproduced as photoengravings and so should be typed carefully and in the *exact* format desired. Authors will be charged for any new engravings necessitated by changes in proof. Use superscript lower-case italic letters (*a*, *b*, *c*) for table footnotes, which should be typed immediately below the table. Type tables at least double-spaced, including titles and footnotes. Do *not* underline table titles; reserve underlining for text that is to be italicized.

Equations should be typewritten and with the number placed in parentheses at the right margin. Reference to equations should use the form "Eq. (3)" or simply (3). Superscripts and subscripts should be typed or handwritten clearly above and below the line, respectively. Use the exponent $^{1/2}$ wherever possible.

"Style." In general, authors should be guided by *A Manual for Authors* ($1.00) published in 1962 (and revised in 1966) by the American Mathematical Society, P.O. Box 6248, Providence, R. I. 02904, **but**

References should be styled and punctuated according to the following examples:

1. G. G. Lorentz, The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239–251. [Do not underline titles of articles or books.]
2. J. T. Scheick, Some problems in approximation theory, Ph. D. Dissertation, Syracuse University, June, 1966.
3. E. Hille and R. Phillips, "Functional Analysis and Semi-groups," American Mathematical Society, Providence, R. I., 1957. [For books, include name of publisher, city, state, and year.]

For unpublished lectures or symposia, include title of the paper, name of the sponsoring society in full, and date. For journal names, follow *Abbreviations of the Names of Scientific Periodicals Reviewed in Mathematical Reviews*, Volume 37, 1969 (reprint $1.00, American Mathematical Society). Cite references in the text by an Arabic number between square brackets, as [1], [1, 2], [1, Theorem 1.5], etc. Type references double-spaced throughout.

Footnotes in text should be avoided if at all possible. If they must be used, identify by superscript numbers and type together on a separate page, double-spaced.

Figures. All illustrations are to be considered as figures. Number each graph or drawing in sequence with Arabic numerals. Supply a descriptive legend for each figure. Type legends double- or triple-spaced consecutively on a separate sheet. **Figures must be submitted in a form suitable for reproduction.** Plan figures to fit the proportion of the printed page. Use a professional lettering set on the original so that the letters and numbers are large enough and "open" enough to take a reduction of 50–60% without filling in with ink. Do not include background grids; however, on paper with blue lines the grid can be eliminated in the process of photoreproduction. Identify each figure in a margin with the name of the journal, author's name, and figure number; avoid marking the backs of figures.

Proofs. Galley proofs will be sent to the author, with a reprint order form. Authors will be charged for alterations in excess of 10% of the cost of composition.

Reprints. Fifty reprints without covers will be provided free of charge. Additional reprints may be purchased; an order form will be included with proofs.

JOURNAL OF ECONOMIC THEORY

Volume 24, Number 1, February 1981

Copyright © 1981 by Academic Press, Inc.
All Rights Reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc. (21 Congress Street, Salem, Massachusetts 01970), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre-1981 articles are the same as those shown for current articles.

Published bimonthly at 41 Tempelhof, Bruges, Belgium by
Academic Press, Inc., 111 Fifth Avenue, New York, N.Y. 10003.

1981: Volumes 24–25. Price $138.00 U.S.A.; $165.00 outside U.S.A.
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publisher. All correspondence and subscription orders should be addressed to the office of the Publishers at 111 Fifth Avenue, New York, N.Y. 10003. Send notices of change of address to the office of the Publishers at least 6 to 8 weeks in advance. Please include both old and new addresses. Postmaster, send changes of address to *Journal of Economic Theory*, 111 Fifth Avenue, New York, New York 10003.

Second class postage paid at Jamaica, N.Y. Air freight and mailing in the U.S.A. by Publications Expediting, Inc., 200 Meacham Avenue, Elmont, New York 11003.

Copyright © 1981 by Academic Press, Inc.
Printed in Bruges, Belgium, by the St. Catherine Press, Ltd.

JOURNAL OF ECONOMIC THEORY 24, 1–39 (1981)

Rawls, Edgeworth, Shapley, Nash: Theories of Distributive Justice Re-examined

MENAHEM E. YAARI*,†

Department of Economics, Hebrew University, Jerusalem, Israel

Received October 4, 1979; revised January 23, 1980

I. INTRODUCTION

While preparing for a seminar on "Economics Justice," I was rather perplexed to find the existing theory of distributive justice in what looked like a state of considerable disarray. Here was an enormous body of writing, dating back almost as far as writing itself, that was as amorphous as putty and almost as yielding when you tried to sink your teeth in it. One could hardly find two writers agreeing on what a third writer had said, to say nothing of the existence of an agreed body of assertions—no matter how small or sketetal—that everyone could adopt as a common point of reference. And it was on this kind of ground that economists were merrily running along, producing such things as "Benthamite" or "Rawlsian" taxation schemes.

Having mentioned John Rawls, I should hasten to say that, in my opinion, one of the most valuable contributions of Rawls' book [19] is indeed its attempt to answer the foregoing criticism. For Rawls had undertaken to state precisely what questions a theory of distributive justice is meant to answer and to lay down a general methodology that would make the answers to these questions deducible from certain ethical premises. Rawls then

---

* This paper is an outgrowth of a seminar that took place at the Hebrew University of Jerusalem in 1977/1978. I had the pleasure of sharing the responsibility for that seminar with Avishai Margalit, and it is here that I wish to record my debt to him, incurred through many conversations both in and out of the classroom. Thanks are due also to Maya Bar-Hillel, to Gideon Schwartz, and to Amartya Sen for their comments.

† In this paper, I have made an attempt to provide a coherent account of the issues that make up the theory of distributive justice, to the best of my understanding. The paper is an economist's way of reading the theory of distributive justice and it focuses on issues which economists tend to raise when discussing problems of distribution. Yet, I should like to emphasize that the paper is intended also for non-economists, and this fact has had an obvious influence upon the style of presentation. To a large extent, this is an essay in the history of thought, so the reader should be alerted not to look for the types of results which usually appear in papers being published in this *Journal.*

0022-0531/81/010001-39$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

proceeded to use this methodology to construct his own theory, called "justice as fairness," and to compare it with other theories of distributive justice, most notably with utilitarianism. Rawls makes a very determined effort to define utilitarianism (as a theory of distributive justice) in terms of the framework that he himself had laid down, so that the differences between the two theories might come into focus. But here I find that Rawls meets with but limited success, and his discussion of utilitarianism, in many instances, is vague and imprecise. Witness the many comments that have appeared since 1971 on the relationship between Rawls and utilitarianism (Arrow [2], Barry [5, Chap. 10], Gordon [9], Harsanyi [12], Lyons [15], Sen [22], among others) no two of which seem to agree on what the issues are, let alone on how they might be resolved. Seeing that this was the case, and with a seminar soon to get underway, I had no choice but to sit down and undertake the exercise of writing precisely what it was, to the best of my understanding, that utilitarianism as a theory of distributive justice did say, and then to do the same with what I understand to be Rawls' position, so as to be able to carry out a rigorous comparison between the two. I started by taking advice from Rawls himself, who says [19, p. 22] that "in order to bring out the underlying differences in the simplest way... the kind of utilitarianism I shall describe is the strict classical doctrine which receives perhaps its clearest and most accessible formulation in Sidgwick." Now, I did not find Sidgwick [24] all that clear, and at certain points I had the feeling that he was being deliberately vague. So, I decided to call in F. Y. Edgeworth as a kind of reinforcement. Edgeworth's [7] articulate and methodical account of utilitarian distribution theory is more orthodox and audacious than Sidgwick's and yet, at the same time, Edgeworth clearly sees himself as building upon the foundations that Sidgwick had laid. Therefore, any contrast that one might draw between Rawls and Sidgwick would no doubt be accentuated and brought more sharply into focus with Edgeworth in the picture.

My version of the confrontation between Sidgwick and Edgeworth on the one hand and Rawls on the other ended very differently from the confrontation that Rawls himself had set up in his book. Using a standard minimax theorem, I was able to argue that, under very general assumptions, the Sidgwick–Edgeworth theory and the Rawlsian theory are equivalent. So, if one still wants to hold the view that Rawls and classical utilitarianism are fundamentally different from each other, then one would have to show either that I have misstated the Sidgwick–Edgeworth position or that I have misstated Rawls' position, or else that the general assumptions under which the two positions are equivalent are fundamentally controversial and that the difference between the two theories rests on the failure of these assumptions.

The Sidgwick–Edgeworth brand of utilitarianism has long been passé, mainly because the pseycho-physical foundations of utility have been debunked and exposed as useless pseychologism. This immediately brings up a very important question: Can a theory of distributive justice that is founded upon utilitarian principles be consistent with modern utility theory? One would have thought that finding the answer to this question would by now be a matter of looking up the appropriate page in some standard text on social philosophy, social choice, or welfare economics. And a very promising source to look up is Chapter 9 of Sen's fine book [21]. But neither this source nor any other source that I have been able to consult gives a definitive answer to the question posed above. The feeling one gets is that, with so many "impossibility theorems" floating around, the answer is probably that a theory of distributive justice which would be both utilitarian and utility–theoretic is impossible. My own position on this question is less pessimistic, and I derive my mild sense of optimism from a framework suggested by Shapley [25] which appears to me to provide a powerful tool for constructing utilitarian theories of social interaction. (To make my position clear, I should emphasize that the reference here is to the general methodology of Shapley's paper, and not to the specific conept—the so-called Shapley value—which he uses as a solution for situations of social choice with conflicting interests.)

It makes no sense to try and carry out a direct comparison between Rawls [19] and Shapley [25] because the two theories rest on different foundations. However, it does make sense so try and see what happens when Shapley's methodology is applied to Rawls' ethics. Such an experiment, one might hope, would lead to a theory of distributive justice that would be true to Rawls' principles and yet be free of conflict with modern utility theory. Intrigued to see what would emerge, I performed the experiment, only to find myself re-discovering Nash's [17] theory of bargaining.

These tales and others will be told in the sequel.

II. FRAMEWORK

I propose to study distributive justice within the general framework of social choice theory. (See, for example, Sen [21].) To capture the essentials of a simple but non-trivial distribution problem, let me start with two building blocks: First, a non-empty set X and, second, a list of n real-valued functions $u_1, u_2, \ldots, u_n$, defined on X. The set X is the universe of discourse and its elements are entities among which society could conceribably be called upon to exercise a choice. For the sake of concreteness, I shall refer to the elements of X as *distribution schemes*, with the interpretation of this term depending on the specific context of any given discussion. Thus, the elements of X could be given such divergent interpretations as alternative constitutions, alternative institutional structures, alternative tax structures, alternative ways of dividing up the net national product, alternative ways of cutting a cake at a bithday party, and so forth. The set X of all conceivable distribution schemes will be referred to as the *choice space*. As for the functions $u_1, u_2, \ldots, u_n$, they will be referred to as the *utilities* of the individuals who make up the society under consideration. In other words, "society" consists of n individuals who are identified by the utilities $u_1, u_2, \ldots, u_n$ all of which are functions assigning real numbers to elements of the choice space X. Utilities are assumed to convery information about the individuals' preferences. The formula

$$u_i(x) > u_i(y)$$

has the obvious interpretation that individual i prefers the scheme x over the scheme y, in the sense that i would rather see x chosen than y. But utilities can also be assumed to convey information about the *itensity* of preferences. For example, the formula

$$u_i(x) - u_i(y) > u_i(w) - u_i(z)$$

could be taken to mean that individual i's preference for x over y is stronger than his preference for w over z. Statements of this form will not appear explicitly in the present essay. However, much of the discussion will be based, implicitly, on the assumption that such intensity statements are meaningful. Now, the most demanding interpretation of the utilities $u_1, u_2, \ldots, u_n$ is one where they are assumed also to convey information about interpersonal comparisons. For example, one could think of the formula

$$u_i(x) > u_j(y)$$

as comparing individual i's position when the scheme x is chosen with individual j's position when the scheme y is chosen, and asserting the former to be more favorable than the latter. The assumption that such interpersonal comparison statements are meaningful is very strong and hardly warranted. Nevertheless, some authors do assume these interpersonal comparisons to be meaningful and therefore, in stating these authors' positions, it will for a while be ecessary to make the same assumption here as well. This issue will be examined in detail in Sections VII and X.

The choice space X consists of all conceivable distribution schemes that might arise as options in any hypothetical social choice situation. Now, every *specific* social choice situation is, however, characterized by a *subset* of X, consisting of those options that happen to be feasible in that choice situation, taking into account such things as resource limitaions, institutional rigidities, etc. This gives rise to the notion of the *set of feasible distribution schemes* (or the *feasible set*, for short) for which the symbol F will be used.

Given the feasible set F, one can go ahead and define a set U in the following manner:

$$U = \{(a_1,\dots,a_n) \mid u_1(x) = a_1, u_2(x) = a_2,\dots, u_n(x) = a_n \text{ for some } x \in F\}.$$

U is called the set of *feasible utility profiles* (or the *utility-possibilities set*). If one defines a vector-valued function $\mathbf{u}$ by writing $\mathbf{u}(x) = (u_1(x),\dots, u_n(x))$ for all $x \in X$, then U is given simply by the statement

$$U = \mathbf{u}(F),$$

U being the image of F under $\mathbf{u}$. Each choice that society can feasibly make is translated into a "profile" (an n-tuple) of utility numbers and the collection of these utility profiles is the set U.

It is now possible to state a fundamental assumption, which I shall take to be an axiom in much of the present analysis (the removal of this assumption will be discussed in Section IX).

PRINCIPLE OF NON-POLARIZATION. *The set U, of feasible utility profiles, is convex.*

The first thing to be noted is that this principle would always be fulfilled if society were allowed to make randomized choices (such as to announce an option under which the toss of a coin would determine which of the two schemes, x or y, will be selected) and if all individuals have von Neumann–Morgenstern utilities. In the absence of randomized choices, the Principle of Non-polarization could still be shown to hold in many specific settings as, for example, in a setting where the object is to split a bundle of perfectly divisible and freely disposable goods among n individuals having concave utilities. In any case, it should be clear that the Principle of Non-polarization can only be defended in a serious way if one is willing to accept the point of view that utilities convey information not only on preferences but also on the *intensity* of preferences. And even with this much granted, it is still easy to construct theories of social interaction that would violate the Principle. (A good example would be a theory of distribution in an economy where production is subject to increasing returns to scale.) My own impression is that such fundamental departures from the Principle of Non-polarization lead, invariably, to social theories that are Marxian in spirit (as is indeed the case when production is subject to increasing returns to scale). For it is basically a Marxist tenet which holds that society is always faced with polarized alternatives—favoring one group or another—with no way for an intermediate compromise. However, for the systems that I shall attempt to discuss in the forthcoming pages, the adoption of the Principle of Non-polarization is certainly consistent, even though the principle may not actually be written down in so many words. (Rawls [19] mentions convexity only in one place—p. 68—while discussing the idea of efficiency. He does not seem to attach much importance to the notion of convexity but, as far as I can see, his analysis is entirely in keeping with this notion. It is clear, for example, in the discussion of the Original Position [19, pp. 130–136] that in Rawls' view, the concept of justice is *prior* to the concept of an institutional structure. Thus, in a very fundamental sense, if society ever finds itself constrained by an institutional structure that violates non-polarization, then it is possible that the structure is itself unjust. This is precisely where a Marxian would rise to object. For him, giving justice a priority over institutional structure would be out of the question.)

The framework for the analysis is now complete in the sense that we have before us a well-defined social choice problem, namely, the problem of how to select a distribution scheme from the set F of feasible schemes, given the utilities $u_1,\dots,u_n$. What I propose to do in the sequel is to examine how, under the Principle of Non-polarization, certain systems of social thought—notably some variants of utilitarianism and Rawls' justice as fairness—would treat this problem and how these treatments are related to one another.

But before I can turn to this task, I find it necessary to enter a brief disclaimer concerning the priority of liberty. I disagree with Rawls' contention that utilitarianism is incapable of granting an absolute priority to liberty. However, this issue is not directly related to the topics that will be discussed in this essay. Therefore, it seems desirable to sidestep this issue altogether by agreeing that the problem shall be limited to the case where the choice which society is called upon to make is solely in matters unrelated to liberty. More precisely, let us assume, for the purposes of the present discussion, that the set F of feasible distribution schemes has the following property. If x and y are elements of F, then x and y lead to identical distributions of liberty among individuals. On the liberty scale, then, all feasible distribution schemes are equal.

III. THE UTILITARIAN PROGRAM

The first thing that a utilitarian—*any* utilitarian—would do when faced with the problem of how society should select from the feasible set F is to construct a measuring rod, in the following manner. For any n-tuple of nonnegative weights, $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_n)$ having the property that $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$, and for any distribution scheme x (feasible or not) the utilitarian would define $W(x, \boldsymbol{\lambda})$ by writing

$$W(x,\boldsymbol{\lambda}) = \lambda_1 u_1(x) + \lambda_2 u_2(x) + \cdots + \lambda_n u_n(x)$$

so that $W(x, \boldsymbol{\lambda})$ is the weighted average, using the set of weights $\boldsymbol{\lambda}$, of the individuals' utilities at x. Let S be the set describing all the ways of assigning weights to individuals:

$$S = \left\{ \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \,\middle|\, \begin{array}{l} \lambda_1, \lambda_2, \ldots, \lambda_n \text{ are non-negative real} \\ \text{numbers and } \lambda_1 + \lambda_2 + \cdots + \lambda_n = 1 \end{array} \right\}.$$

(S is usually referred to as the *unit simplex* of dimension $n - 1$.) Then, W is a real-valued function defined on the Cartesian product $X \times S$. This function will be called *the utilitarian form.*

Now, why should this function, W, be the first thing that a utilitarian (and as we shall see, not only a utilitarian) would consider? The answer is given by the Pareo Principle, which may be stated as follows:

THE PARETO PRINCIPLE. *Let x be a feasible distribution scheme* $(x \in F)$. *If there exists a feasible scheme y* $(y \in F)$ *satisfying $u_i(y) > u_i(x)$ for $i = 1, \ldots, n$, then the scheme x is irrelevant for social choice.*

Under this principle, only so-called *Pareto-efficient* distribution schemes are relevant for social choice, provided of course that such Pareto-efficient schemes exist in sufficient abundance. (A feasible scheme $x \in F$ is said to be Pareto-efficient—in the weak sense—if there does *not* exist a feasible scheme $y \in F$ with $u_i(y) > u_i(x)$ for $i = 1, \ldots, n$.)

The connection between the Pareto Principle and the utilitarian form W emerges as soon as one writes down the following well-known proposition:

PROPOSITION. *Under the principle of non-polarization, a feasible scheme x* $(x \in F)$ *is pareto-efficient if and only if there exists a system of weight $\boldsymbol{\lambda}$* $(\boldsymbol{\lambda} \in S)$ *such that $W(x, \boldsymbol{\lambda}) \geqslant W(y, \boldsymbol{\lambda})$ for all feasible schemes y* $(y \in F)$.

In other words, a feasible scheme x is Pareto-efficient if and only if it *maximizes* the utilitarian form W, for some choice of weights $\boldsymbol{\lambda}$. The Pareto Principle may therefore be re-phrased as follows. If Pareto-efficient schemes can be assumed to exist then, under non-polarization, the problem of selecting a scheme x from the feasible set F is equivalent (up to a matter of indifference for all individuals) to the problem of choosing a set of weights $\boldsymbol{\lambda}$ in S and then looking only at those schemes x for which $W(x, \boldsymbol{\lambda})$ is a maximum on F. This makes the utilitarian form W relevant not only in utilitarian theories, but in any theory which adopts the Pareto Principle.

The Pareto Principle enjoys wide acceptance, by utilitarians and by non-utilitarians alike (see, e.g., Rawls [19, p. 17]). The main reason for accepting this principle in a theory of distributive justice is that a failure to do so would lead to the untenable position that justice may have to be enforced, against the will of *all* individuals. It is one thing to say that justice may require enforcement upon *some* individuals, perhaps even as many as all except one. But if all individuals hold, unanimously, that a distribution scheme, say x, is strictly preferable to some other scheme, say y, then it is very hard to acept that justice should demand, nevertheless, that y be chosen over x. This point has been brought forth very effectively by Nozick's [18, pp. 160–164] famous "Wilt Chamberlain" example. (Nozick himself draws a very different moral from this example, namely, that "no end-state principle or distributional patterned principle of justice can be continuously realized without continuous interference with people's lives" [18, p. 163]. I confess that I am puzzled by this conclusion taking the place of what appears to me to be the obvious conclusion, namely, that one of the principles of distributive justice had better be the Pareto Principle. Nozick's own explanation [18, p. 164, footnote] does little to relieve my sense of puzzlement, because his new example—unilateral transfers or bequests—is still based on taking an initial situation which is not Pareto-efficient.) Now, strong and persuasive as the case for the Pareto Principle might be, there is of course no guarantee against another principle, perhaps equally as persuasive, coming up later in the discussion and turning out to be incompatible with the Pareto Principle. In the face of such a conflict of principles, some people would choose to abandon one principle, others would abandon another, and still others would abandon the analysis altogether. (See McClelland and Rohrbaugh [16] for an attempt to study the extent to which people who are forced to make judgements in the face of conflicting principles tend in fact to abandon the Pareto Principle.) In the present essay, no postulates will arise that are in conflict with the Pareto Principle, so adopting it here seems to be in order.

Convinced of the reasonableness of the Pareto Principle, the utilitarian is now ready to lay down a general procedure for social choice. I shall refer to this procedure as "the utilitarian program":

THE UTILITARIAN PROGRAM. *First, use ethical judgments to determine what the right set of weights, say λ^*, should be, for the given feasible set F. Then select the distribution scheme x^* in such a way that $W(x, \lambda^*)$ shall be a maximum, over all x in F, at $x = x^*$.*

At this stage, I neglect the problem of ties and how they should be broken.

IV. ON MAXIMIZING THE SUM-OF-UTILITIES: A CRITICAL REMARK

As you read the literature on social choice (and on distributive justice as a branch of social choice theory) you find a tendency among authors to identify utilitarianism with the proposition that society should always select the alternative that maximizes the sum of the individuals' utilities. I confess that I find the cavalier manner in which this pronouncement is usually made rather infuriating. Even as careful a writer as Sen [22, 23] is, in my opnion, not careful enough on this point.

Equating utilitarianism with the maximization of the sum-of-utilities is equating it with the proposition that the "correct" set of weights, $\boldsymbol{\lambda}^*$, to be used in the utilitarian form W is the same for *all* feasible sets F. This point of view sees the utilitarian as solving social choice problems by looking always for a Pareto-efficient scheme at which the rates of transfer of utility among individuals are in accordance with some universal constants $\boldsymbol{\lambda}^*$. Even the most naive utilitarians cannot be characterized as having argued in this way.

In one (trivial) sense, the assertion that utilitarianism coincides with the maximization of the sum-of-utilities is always true, by definition. Consider the utilitarian program, described in the last paragraph of Section III and suppose that, in one way or another, a set of weights $\boldsymbol{\lambda}^* = (\lambda_1^*,\ldots,\lambda_n^*)$ has already been determined. Let the functions $v_1, v_2,\ldots, v_n$ be defined by $v_1(x) = \lambda_1^* u_1(x)$, $v_2(x) = \lambda_2^* u_2(x),\ldots, v_n(x) = \lambda_n^* u_n(x)$ for all $x \in X$, where $u_1,\ldots, u_n$ are the individuals, utilities, given originally. Then, by this definition, the utilitarian program would always have x^* chosen to maximize the sum $v_1(x) + v_2(x) + \cdots + v_n(x)$. Clearly, on this interpretation, to equate utilitarianism with maximization of the sum-of-utilities is to confuse *ex-ante* with *ex-post*, because the utilities whose sum is being maximized are themselves an outcome of the analysis.

But what other interpreation can there be for the sum-of-utilities maxim? Arguing in a somewhat simple-minded fashion, one can say that the very act of writing the symbol "+" in an expression like $u_1(x) + u_2(x) + \cdots + u_n(x)$ means that the utilities $u_1, u_2,\ldots, u_n$ have been jointly calibrated and converted into a *common* unit, so that the u_i's whose sum is being taken are, in fact, the v_i's of the previous paragraph. There is no way to calibrate the utilities $u_1,\ldots, u_n$ and convert them to a common unit, except by using conversion factors which, even if equal in magnitude, are not equal in unit of measurement. Thus, by forcing the addition of utilities to be well-defined, one automatically introduces the assumption that utilities have been re-defined so as to incorporate the appropriate conversion rates. This looks almost like a *proof* that the (trivial) *ex-post* interpretation of the sum-of-utilities maxim is the only one possible. Almost, but not quite. After all, there do exist structures that are amenable to extensive-additive measurement. Let us, for a moment, take a look at the conditions that would have to be satisfied in order for the utilities of different individuals, jointly, to be a structure admitting extensive-additive measurement. There are four such conditions (here, I follow Krantz, Luce, Suppes, and Tversky [13, pp. 71–75]):

(i) There exists a non-empty set, E, of unambigously defined entities having the attributes of pleasure and pain. Elements of E will be referred to as *experiences.*

(ii) There exists a binary relation $\gtrsim$ on E which is reflexive, transitive, and connected. (That is, $\gtrsim$ is a complete weak order of E.) If e and e' are experiences, then the statement $e \gtrsim e'$ has the interpretation that e is at least as pleasurable or at most as painful as e'.

(iii) There exists a binary operation $*$ on E, called *concatenation* (or *joining together*) so that any two experiences (not necessarily of the same individual) may be concatenated, and the result is once again an experience. This binary operation satisfies such algebraic laws as associativity and cancellation.

(iv) The relation $\gtrsim$ and the operation $*$ fit together nicely in two different ways. First, if e, e', f, f', are experiences such that $e \gtrsim f$ and $e' \gtrsim f'$, then $e * e' \gtrsim f * f'$. Second, experiences are Archimedean, in the sense that if ε is any experience such that $\varepsilon * \varepsilon \neq \varepsilon$ then, by concatenating ε with itself sufficiently many times, one gets a new experience lying outside any preassigned $\gtrsim$-interval of the form $\{e \in E \mid e_1 \gtrsim e \gtrsim e_2\}$.

Anyone equating utilitarianism with the sum-of-utilities maxim (in a nontrivial *ex-ante* sense) is actually equating utilitarianism with the adoption of these four conditions, (i) through (iv), as postulates.

Technical remark. Some readers will object here and say that in order to derive the sum-of-utilities maxim it is enough to have a social preference order, defined on utility profiles, which is symmetric and additively separable. But this would be wrong, because utility profiles are not invariant to changes of individuals' utility indicators.

As far as I know, the only utilitarian who could conceivably be stretched to the point of adopting these four postulates is Edgeworth, who thought that the notion of a "just-perceptible increment" could be used to construct a universal unit for measuring pleasures and pains (see Section VI). But as we shall see, even Edgeworth's theory of social choice has more to it than a mere maximization of the sum-of-utilities.

Now, what about the argument that derives the sum-of-utilities maxim from considerations of choice under uncertainty? The argument goes something like this (see Harsanyi [11]): Suppose you had to select a distribution scheme x from the feasible set F before knowing what individual characteristics you were going to have. In particular, suppose you believed that your utility could be one of the functions $u_1, u_2, \ldots, u_n$ and that it would be just as likely to be one or the other. Then, if your choice behavior satisfies certain axioms, you would select x^* to maximize the expected utility

$$\frac{1}{n} u_1(x) + \frac{1}{n} u_2(x) + \cdots + \frac{1}{n} u_n(x)$$

over all x in F, which is the same thing as maximizing the sum $u_1(x) + u_2(x) + \cdots + u_n(x)$. This reasoning is valid for all participating individuals, and therefore, by unanimity, social choice should seek to maximize the sum-of-utilities after all. How about that? Well, to begin with, I should hasten to set the record straight by noting that this is not what Harsanyi proves. Indeed, Harsanyi [11, Theorem 5] is very precise in stating that the rule which he seeks to establish using these probabilistic considerations is one where social choice is governed by the maximization of a *weighted* sum-of-utilities, with weights that are not necessarily equal and have nothing to do with the probabilities of the underlying decision-making situation. As for the point itself, one must remember only that probabilities, like relative frequencies, are pure numbers (unit-less magnitudes). Therefore, if $p_1, p_2, \ldots, p_n$ are probabilities, then the expected utility $p_1 u_1(x) + p_2 u_2(x) + \cdots + p_n u_n(x)$ is well-defined only after the functions $u_1, u_2, \ldots, u_n$ are jointly calibrated and converted to a common unit. We are right back where we started.

A final comment may be in order, for readers who might wonder how summing utilities can meet with so much wrath when, all along, *weights* were being added together, in expressions like $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$. Are not λ_i and λ_j measured in different units? Is writing $\lambda_i + \lambda_j$ not like summing apples and oranges? The answer should be apparent upon a moment's reflection: In the maximization of the utilitarian form, $W(x, \boldsymbol{\lambda})$, only the *relative* magnitudes of the weights (i.e., ratios of the form λ_i/λ_j) play a role. Therefore, each weight may be divided by the sum of all weights, without changing anything.

V. MODERN UTILITARIANISM: SHAPLEY'S PROPOSAL

Anyone proposing a theory of social interaction which is consistent with the Principle of Non-polarization and with the Pareto Principle can be thought of as carrying out what I have termed "the utilitarian program." Yet it seems that for a theory of social interaction to be called *utilitarian*, it must also, in one way of another, incorporate the idea of *unility trade-offs* among individuals. When somebody proposes a social change that would benefit me but harm you, it becomes the utilitarian's task to offer a way of weighing my benefit against your loss. For the modern utilitarian, this task is very difficult indeed, because he must somehow subsist on what is left after modern utility theory and Arrow's impossibility theorem have each taken their shares. A general and ingenious proposal for meeting this modern utilitarian challenge has been made in a contribution due to Shapley [25]. Shapley's proposal has already been applied in several directions, and the results so far have been encouraging. (See, for example, Aumann and Kurz [4] who use Shapley's procedure to study the effect of the political structure on taxation.)

What Shapley says is this. Consider the weights $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ appearing in the utilitarian form $W(x, \boldsymbol{\lambda})$. These weights play the role of conversion rates that transform individual utilities into some universal unit. In particular, these weights act also as *rates of transformation* or *rates of exchange* between individual utilities. For example, if society has chosen the weights $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ then, by so doing, society has in fact decreed that one unit of individual i's utility is exchangeable for λ_j/λ_i units of individual j's utility. In other words, every choice of weights $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ defines a system of transferable utility between individuals (provided none of the λ_i vanishes) where the ratio λ_j/λ_i defines the rate at which utility side-payments between individuals i and j are to be made. But once utility is transferable, so says Shapley, we can bring to bear a very powerful theory, namely, the theory of games with transferable utility. So, for every choice of weights, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$, we are told to construct a game, say $G(\boldsymbol{\lambda})$, which would describe our social choice problem in terms of what individuals can achieve under each distribution scheme, on the assumption that the total utility obtained by any group of individuals can be divided up in any manner among its members, using the rates of transfer $\lambda_1, \ldots, \lambda_n$. There are usually several ways to construct the game $G(\boldsymbol{\lambda})$, with different sets of rules for how the game is to be played. This fact accounts for the generality of Shapley's proposal. Now, once the game $G(\boldsymbol{\lambda})$ is constructed, one can appeal to some solution concept, for games with transferable utility, to tell us how society should divide up the total utility which it can obtain under the rules of the game. (The solution concept recommended by Shapley is the so-called "modified Shapley value" but this is not essential to his proposal.) Let us say, then, that some solution concept has been selected and that, under this solution concept, the distribution of total utility among individuals in the game $G(\boldsymbol{\lambda})$ is given by $\boldsymbol{\mu}(\boldsymbol{\lambda}) = (\mu_1(\boldsymbol{\lambda}), \ldots, \mu_n(\boldsymbol{\lambda}))$. That is, for each set of weights $\boldsymbol{\lambda}$, the game $G(\boldsymbol{\lambda})$ is constructed and then resolved into an n-tuple of numbers, $(\mu_1(\boldsymbol{\lambda}), \ldots, \mu_n(\boldsymbol{\lambda}))$, describing the utility payoffs that individuals get in the game $G(\boldsymbol{\lambda})$, under the adopted solution concept. In general, the utility profile $(\mu_1(\boldsymbol{\lambda}), \ldots, \mu_n(\boldsymbol{\lambda}))$ *will not be feasible* in the original distribution problem. That is, there will not in general exist a feasible distribution scheme x $(x \in F)$ having the property that $u_1(x) = \mu_1(\boldsymbol{\lambda})$, $u_2(x) = \mu_2(\boldsymbol{\lambda}), \ldots$, $u_n(x) = \mu_n(\boldsymbol{\lambda})$. If this is the case—the utility profile $(\mu_1(\boldsymbol{\lambda}), \ldots, \mu_n(\boldsymbol{\lambda}))$ not being feasible—then the system of weights $\boldsymbol{\lambda}$ has *failed to justify itself* in the sense that these rates of utility transfers, $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$, do not correctly reflect the realities of the underlying choice situation. Therefore, a new set of rates for utility transfer must be examined and the process must be repeated until a set of weights, say $\boldsymbol{\lambda}^* = (\lambda_1^*,..., \lambda_n^*)$, is found having the property that the utility profile $(\mu_1(\boldsymbol{\lambda}^*),..., \mu_n(\boldsymbol{\lambda}^*))$ is feasible in the original distribution problem. Only a set of weights $\boldsymbol{\lambda}^*$ having this property can be a candidate for use in the utilitarian form $W(x, \boldsymbol{\lambda})$.

For the sake of brevity, let me say that a set of weights $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$, in the unit simplex S, is *self-justifying* if there exists a feasible distribution scheme x $(x \in F)$ such that $u_1(x) = \mu_1(\boldsymbol{\lambda})$, $u_2(x) = \mu_2(\boldsymbol{\lambda}),..., u_n(x) = \mu_n(\boldsymbol{\lambda})$, where $(\mu_1(\boldsymbol{\lambda}),..., \mu_n(\boldsymbol{\lambda}))$ is the distribution of utilities in the solution of the game $G(\boldsymbol{\lambda})$. Shapley's contribution consists of two separate items. First, he proves that, for a large class of games $G(\boldsymbol{\lambda})$ and using any solution concept satisfying certain general requirements (continuity and the Pareto Principle), *there always exists in S a set of weights $\boldsymbol{\lambda}^*$ which is self-justifying.* Second, Shapley offers a heuristic argument (based on "counting equations and unknowns") to show that, in most cases, there exists in S *one and only one* set of weights $\boldsymbol{\lambda}^*$ that is self-justifying. (Here, the meaning of "in most cases" is vague, but it can be made precise.) Now, if it happens in any specific model (as, for example, in Aumann and Kurz [4]) that self-justifying weights are in fact unique, then the argument for using precisely these weights in the utilitarian form W is very hard to resist. For suppose that society decides to select some Pareto-efficient distribution scheme, say x^*, from the set of feasible schemes F. Under non-polarization, there will exist a set of weights $\boldsymbol{\lambda}^* = (\lambda_1^*,..., \lambda_n^*)$ corresponding to x^*. Thus, by selecting the scheme x^*, society is in fact asserting that it is correct to pretend that utility is transferable between individuals at the rates $\lambda_1^*,..., \lambda_n^*$. But if pretending that utility is transferable at the rates $\lambda_1^*,..., \lambda_n^*$ does not lead back to a distribution scheme where these are, in fact, the local rates of transfer (i.e., back to x^* itself) then the pretending was clearly incorrect. Thus, Shapley's proposal may be summarized as follows: *Never select a distribution scheme whose implicit rates of utility transfer would lead you to want to move away from that scheme.* This sounds like a rather minimal consistency requirement, but in many cases it suffices to determine a unique selection.

But how would ethical principles, without which there cannot be a theory of justice, fit in Shapley's framework? The answer, quite clearly, is that such principles would have to go in though the specification of the games $G(\boldsymbol{\lambda})$ and through the choice of a solution concept for these games. For example, it is possible to construct the games $G(\boldsymbol{\lambda})$ in such a way as to incorporate the principle that each person's utility payoff should reflect the share of total wealth that this person brings into the game at the outset (something like "to each according to his wealth"). And there is a specific way to do this, whereupon Shapley's procedure leads directly to the distributive mechanism known in economics as the competitive equilibrium. In principle, one can design the games $G(\boldsymbol{\lambda})$ to reflect any ethical or distributional postulate and then use Shapley's procedure to derive the corresponding distributive mechanism. In Section X, I shall attempt to do just that, using Rawls' Difference Principle as the underlying ethical postulate.

VI. CLASSICAL UTILITARIANISM

My task in this section is to try and translate the classical utilitarian doctrine into the language of finding an appropriate set of weights, $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_n)$, for the utilitarian form $W(x, \boldsymbol{\lambda})$, and picking a distribution scheme x^* that maximizes $W(x, \boldsymbol{\lambda})$ among all feasible schemes x $(x \in F)$. Since classical utilitarians did not all speak in one voice, my translation is bound to contain an element of interpretation and I shall try, needless to say, to convince the reader that may interpretation is indeed correct. But, perhaps more importantly, I shall argue that my interpretation of classical utilitarianism is no different from Rawls' interpretation of the same doctrine. Why this should be important will be seen later, when I discuss the alleged contrast between classical utilitarianism and Rawls' own theory. But first, let me try to formulate as rigorous a statement as I can of classical utilitarianism. In this attempt, I shall rely for the most part on the writings of Sidgwick [24] and Edgeworth [7]. Sidgwick's account of this doctrine is, perhaps, the most balanced and deliberate (sometimes to the point of being noncommittal) while Edgeworth's is propably the most spirited and enthusiastic.

In Section IV, the problem of selecting a set of weights, $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_n)$, for the utilitarian form $W(x, \boldsymbol{\lambda})$ was discussed in terms of the theory of measurement: Can individual utilities be converted into a common unit and, if so, how? For classical utilitarians, this problem was no different from the problem of converting, say, the lengths or weights of individual physical objects into a common unit. Happiness, or pleasure, was thought of as a psycho-physical quantity given to objective measurement that makes it commensurable between individuls (Edgeworth [7, p. 59]). Furthermore, it was proposed that the Weber–Fechner notion of a just-perceptible stimulus could very well be used to construct a universal unit of utility that would enjoy the same degree of efficacy as do the gram or the centimeter in physical science. Thus, Edgeworth says [7, p. 60]: "Just perceivable increments of pleasure, of all pleasures for all persons, are equateable." (See also Edgeworth's discussion of "hedonimetry" [7, pp. 98–102] where he argues that there is no reason for the margin of error in hedonimetric measurements to exceed the margin of error in physical measurements. Sidgwick's position on this [24, 3rd ed., pp. 146–147] is similar to Edgeworth's, albeit more hesitant.)

For classical utilitarians, then, the individual utilities $u_1, u_2, \ldots, u_n$ may be taken, at the outset, as measured in terms of a common unit. End of story? Not quite. The weights $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ still have a major role to play, but they are not interpreted as rates of conversion in a stystem of measurement. Rather, the interpretation of these weights is that λ_i stands for the *proportion of individuals of type i in the population.* In classical utilitarianism, the structure as well as the size of the population was very much an object for social choice, especially after Malthus and his followers had spoken of what would happen if population were not made an object for conscious social policy. Both Sidgwick and Edgeworth are quire clear on this. Sidgwick says [24, 3rd ed., p. 412]: "But a further question arises when we consider that we can to some extent influence the number of future human (or sentient) beings. We have to ask how, on Utilitarian principles, this influence is to be exercised." And Edgeworth [7, p. 56] opens the section on Utilitarian Calculus as follows: "*Problem.* To find (α) the distribution of means and (β) of labour, the (γ) quality and (δ) number of population, so that there may be the greatest possible happiness." Edgeworth speaks expliitly of two different ways for influencing population structure, namely, migration and something called evolution. Migration is clear enough, as is its position among the prominent issues of social policy. As for "evolution," what Edgeworth had in mind was a process affecting population structure and lying somewhere between outright population control and a "Darwinian" theory in which the distribution of happiness is viewed as affecting, possibly, the rates of reproduction and survival.

To make a long story short, we are now given a list of utilities $u_1, \ldots, u_n$, each representing a *type* of individual or a social class (or a "section" as Edgeworth calls it) and the problem of social choice is to select a distribution scheme x and a set of population parameters $k_1, \ldots, k_n$ so as to maximize the quantity

$$k_1 u_1(x) + k_2 u_2(x) + \cdots + k_n u_n(x),$$

among all $(x, k_1, k_2, \ldots, k_n)$ lying in some feasibility set. Here, k_i is the number of individuals of type i, or the size of the ith social class. If one lets the size of the total population be denoted N (that is, $N = k_1 + k_2 + \cdots + k_n$) and if one defines the proportions $\lambda_1, \ldots, \lambda_n$ by writing $\lambda_i = k_i/N$, then the social choice problem becomes one of finding a distribution scheme x, a system of weights $\boldsymbol{\lambda}$, and a population size N so as to maximize—under suitable constraints—the expression $NW(x, \boldsymbol{\lambda})$, where $W(x, \boldsymbol{\lambda})$ is now interpreted as the (weighted) *average happiness* which would result if the policy x

were adopted and if $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_n)$ were the proportions of the various types of people in the population.

All this is taken right out of Sidgwick [24, 3rd ed., p. 413] who says "the point up to which, on Utilitarian Principles, population ought to be encouraged to increase, is not that at which average happiness is the greatest possible, but that at which the product formed by multiplying the number of persons living into the amount of average happiness [i.e., the product formed by multiplying N into $W(x, \boldsymbol{\lambda})$] reaches its maximum." Edgeworth, for this part, refers explicitly [8, p. 19] to this passage in Sidgwick, calling it "the nearest approach to an absolutely new idea of first-rate importance in morals."

The maximization of something like $NW(x, \boldsymbol{\lambda})$ seems to provide a reasonable formulation for Hutcheson's celebrated maxim: "that action is best which procures the greatest happiness for the greatest numbers." $W(x, \boldsymbol{\lambda})$ is "happiness" while N is "numbers" so maximizing their product is an attractive explication of Hutcheson's maxim. It may therefore appear curious that Edgeworth should have been so contemptuous of it. He says [7, p. 117]: "is this [maxim] more intelligible than 'greatest illumination with the greatest number of lamps'?" If one takes the view, as Edgeworth did, that man is a pleasure machine, then by definition, the greater the numbers, the greater the aggregate of pleasures. So, what sense is there in speaking of the greatest pleasure for the greatest numbers? The resolution of this apparent difficulty emerges as soon as the maximization of $NW(x, \boldsymbol{\lambda})$ is spelled out more carefully. For this maximization takes place under *two* separate sets of constraints. One set of contraints is physiological, describing what Sidgwick and Edgeworth call "the capacity for happiness" of the various types of individuals. Specifically, if k_i is the number of individuals of type i and if u_i is the utility level enjoyed by a person of type i, then, in maximizing the sum $k_1 u_1 + k_2 u_2 + \cdots + k_n u_n$, the utility profile $(u_1, ..., u_n)$ must be taken from some specific set, U, which describes the different capacities for happiness of the different types of individuals. For example, if there are only two types of individuals, type 1 and type 2, and if individuals of type 1 have precisely twice the capacity for happiness as do individuals of type 2, then the set U will consist of all pairs (u_1, u_2) of nonnegative real numbers satisfying $u_1 + 2u_2 \leqslant K$, where K is some constant that may be normalized so as to have $K = 1$. In general, the set U describes the psycho-physically determined rates of transformation between the happiness of different types of individuals and it has nothing whatsoever to do with the feasibility or lack of feasibility of one social action or another. (Recall that, in previous sections of this essay, the set U was the image, in utilities-space, of the set F of socially feasible distribution schemes.) By Fechner's law of diminishing ability to perceive incremental stimuli of constant magnitude, U is some convex subset of Euclidean n-space, so that, in this theory, the Principle of

Non-polarization follows from the laws of pseycho-physiology. At any rate, the maximization of $k_1 u_1 + k_2 u_2 + \cdots + k_n u_n$ over all $(u_1, u_2, ..., u_n)$ lying in U is meaningless, since "numbers"—that is, the k_i's— can be made as large as we please, thereby making the maximand arbitrarily large. This is the meaning of Edgeworth's mocking paraphrase—"greatest illumination with the greatest number of lamps."

The maximization problem becomes meaningful when a second family of constraints is recognized, having to do with the availability of resources. If resources were unlimited, then the quantity $NW(x, \boldsymbol{\lambda})$ would have no maximum in the sense that infinitely many people would be able to enjoy infinite happiness. So, the problem of social choice arises only when resources are limited. Sidgwick and Edgeworth themselves did not use the term "resources" but preferred a term that is more general and much more vague, namely, "the means of happiness." Edgeworth discussess this vagueness explicitly [7, p. 69], indicating how difficult it would be to tell how much happiness could be produced from a given amount of means (= resources). It is clear, however, that at any given moment, there exists a *technology* that describes the various ways for converting means of happiness into happiness, and it is through this technology that the limited availability of resources acts as a constraint on social choice. It is quite easy to write all this down very formally, but this seems to me rather an idle exercise. Instead, I shall try to bring out the essentials of the problem in a simpler setting, where total happiness itself acts as the limiting factor. In other words, I shall assume that society must act under the constraint that total welfare (= total happiness = Sidgwick's "quantum of happiness" = Edgeworth's "lot of happiness") cannot exceed some number, say $\bar{W}$. Herc, the reader will no doubt wonder: If $NW(x, \boldsymbol{\lambda})$ cannot exceed $\bar{W}$ then, obviously, one should set $NW(x, \boldsymbol{\lambda}) = \bar{W}$, so what on earth is left to be maximized? The "physiological" requirement that the utility profile must lie in the set U is of no help because, as has already been mentioned, the number N can always be adjusted to have $NW(x, \boldsymbol{\lambda}) = \bar{W}$. Indeed, for *any* "physiologically" feasible pair $(x, \boldsymbol{\lambda})$ one can go ahead and calculate $W(x, \boldsymbol{\lambda})$ and then set N to satisfy $NW(x, \boldsymbol{\lambda}) = \bar{W}$, so that any pair $(x, \boldsymbol{\lambda})$ is as good as any other. It appears as though we keep coming back to Edgeworth's "greatest illumination with the greatest number of lamps" except that now we set the number of lamps in accordance with some fixed total amount of available illumination. But a careful writing down of who-maximizes-what-with-respect-to-which-variables shows that this is not so. Suppose that the weights $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_n)$ have already been determined somehow. Then, there is no question but that the classical utilitarian would select x^* to maximize $W(x, \lambda)$ among all feasible x—i.e., among all "physilogically" feasible utility profiles—and then let total population be determined from the equation $N = \bar{W}/W(x^*, \boldsymbol{\lambda})$. (The set F of feasible distribution schemes plays no role in this discussion but I am retaining it nevertheless, along with phrases like "picking x^* to maximize $W(x, \boldsymbol{\lambda})$ among all feasible x," for the purpose of keeping consistent terminology throughout the essay. It may be worthwhile to point out that, also in the other sections of this essay, it is not the set F as such which determines the structure of a distribution problem, but rather the set U of feasible utility profiles. Here, this same set, U, is given the "physiological" interpretation described above.) Furthermore, society—being fully rational—*knows* that after it decides on the population proportions $\boldsymbol{\lambda}$, it will in fact follow the procedure of selecting x^* to maximize $W(x, \boldsymbol{\lambda})$. Therefore, in deciding which proportions $\boldsymbol{\lambda}$ to pick, society takes this knowledge into account. Now, $W(x^*, \boldsymbol{\lambda})$ will be the average happiness enjoyed by a representative individual, and this average is at a maximum, given the proportions $\boldsymbol{\lambda}$. Let us write

$$x^* = \xi(\boldsymbol{\lambda})$$

to mean that x^* is that choice of x, among all x in F, for which $W(x, \boldsymbol{\lambda})$ is at a maximum. It is assumed that this choice is unique or, more precisely, that other principles will have to be relied upon in case of a tie that needs to be broken. Now, when it comes to picking the proportions $\boldsymbol{\lambda}$ themselves, the object should be to let the greatest possible numbers enjoy the level of happiness $W(\xi(\lambda), \boldsymbol{\lambda})$, this being the maximum happiness enjoyed by a representative individual, under the proportions $\boldsymbol{\lambda}$. In other words, the population proportions should be picked so as to *maximize* N, subject to the constraint that $NW(\xi(\boldsymbol{\lambda}), \boldsymbol{\lambda})$ cannot exceed the total amount of happiness, $\bar{W}$. In symbols, one gets—

$$\max_{\boldsymbol{\lambda} \in S} N \qquad \text{subject to} \quad NW(\xi(\boldsymbol{\lambda}), \boldsymbol{\lambda}) = \bar{W},$$

or, substituting for $\xi(\boldsymbol{\lambda})$—

$$\max_{\boldsymbol{\lambda} \in S} N \qquad \text{subject to} \quad N[\max_{x \in F} W(x, \boldsymbol{\lambda})] = \bar{W},$$

which I now take to be the explication of Hutcheson's maxim. The maximum inside brackets stands for "the greatest happiness" and the maximum of N on the left stands for "the greatest numbers." Each maximization takes place with respect to a different decision variable: The former takes place with respect to the choice of a distribution scheme (or utility profile) and the latter with respect to the choice of a population structure. Edgeworth [7, pp. 61–75] recognizes these two levels of social choice quite explicitly and he offers optimal rules for both of them. His solution for the optimal population structure [7. pp. 70–71] is given by a kind of "bang-bang" policy which, no doubt, sounds rather weird: All persons whose capacity for happiness is greater than some critical value should be encouraged to reproduce at the maximum possible rate, while all persons whose capacity for happiness lies below that critical value should not be allowed to reproduce at all. Edgeworth himself is concerned that this conclusion might "appear outlandish" [7, p. 71] so he hastens to qualify it in various ways, some of them not unamusing. In Section VIII, I shall argue that Edgeworth's fear to appear outlandish was without foundation. For Edgeworth's population policy— so I shall argue—is in fact an early forerunner of Rawls' Difference Principle. The outlandish aspect is not in the principle itself but rather in its application to the problem of population planning. Indeed, if Rawls had tied his Difference Principle to population planning (something like "only those in the least-favored position should be allowed to reproduce") would this not appear equally outlandish?

To recapitulate, what I have been trying to do in these last few pages is to convince the reader that classical utilitarianism, as a framework for social decision making, can be summarized as follows. Find a distribution scheme x^* and a system of weights λ^* so that the pair (x^*, λ^*) shall provide a solution for the problem

$$\max_{\lambda \in S} N \qquad \text{subject to } N[\max_{x \in F} W(x, \lambda)] = \bar{W},$$

where the sets S and F are as defined above. A moment's reflection makes it clear that this optimization problem can be reformulated in a way that avoids explicit reference to N and $\bar{W}$. This leads to another statement of classical utilitarianism, equivalent to the previous one, namely:

CHOICE UNDER CLASSICAL UTILITARIANISM. *Find a distribution scheme x^* and a system of weights λ^* so that the pair (x^*, λ^*) shall provide a solution for the problem*

$$\min_{\lambda \in S} \max_{x \in F} W(x, \lambda).$$

Except for tie-breaking, we have here a complete proposal for how to carry out the utilitarian program (see Section III) telling us which weights, λ^*, to use in the utilitarian form, whose maximization then leads to the choice of a distribution scheme, x^*. The specific interpretation which classical utilitarians had given the weights λ^*—the proportions of the various types of individuals in the population—is completely independent of the formal principle itself, and it is possible to adopt the principle without adopting the interpretation. Thinking of the weights λ^* as population parameters is a consequence of the classical utilitarians' commitment to a pseycho-physiological calculus of pleasures and pains. For anyone not committed to such a calculus, the need to interpret λ^* in this way disappears.

But, interpretations aside, it is very important to note how the weights λ^* are to be determined, according to classical utilitarianism. The guiding principle is clearly one of economy, calling for the weights to be selected in such a way as to stretch a given amount of total social welfare as far as possible. Suppose that, upon working out the classical utilitarians' proposal, we were to find that $\lambda_i^* > \lambda_j^*$, for some i and j: Individual i is assigned greater weight than individual j. Is this because i is more deserving than j? Or, perhaps, because i's needs are greater than j's? Or is there a legal case showing that i is entitled to more than j? Or has i's utility been stated in the wrong units, relative to j's? It is none of these. Individual i is assigned greater weight than individual j merely because such an assignment allows a given quantity of happiness to be stretched farther than would an alternative assignment. Persons are distinguished from one another only according to considerations of efficiency. Ethics, like economics, is seen as derived from a general mechanical principle—say Hamilton's Principle of Least Action—which says that rules of human conduct and behavior, like equations of motion, have their roots in a universal tendency toward efficiency. At this point, how can I resist reproducing that much-cited passage from Edgeworth [7, p. 21]: " 'Mécanique Sociale' may one day take her place along with 'Mécanique Celeste,' throned each upon the double sided height of one maximum principle." This celebrated passage reinforces my belief that the formulation of classical utilitarianism given above is indeed correct. For if one takes the utilitarian form W as the social analogue of the Hamiltonian form and the weights λ as the social analogue of Hamilton's "auxiliary variables" (which always enter linearly) one finds that the analogue of Hamilton's Principle comes out to be precisely that $W(x, \lambda)$ shall be a maximum with respect to x and a minimum with respect to λ.

The impersonal, mechanistic aspect of classical utilitarianism is indeed striking. This is also the conclusion reached by Rawls [19] who examines classical utilitarianism at some length. (In Rawls' discussion, a great deal of emphasis is given to the notion of an impartial sympathetic spectator as being central in classical utilitarianism. This construction—the impartial spectator—had been invented by early utilitarians for the purpose of handling the problem of how individual desires and wants are to be assessed and compared. However, apparently due to the advant of Fechnerian psychophysics, this construction was no longer deemed necessary, so it was simply dropped. This is evidenced by the fact that neither in Sidgwick's account of utilitarianism nor in Edgeworth's can an appeal be found to the notion of an impartial spectator. In any event, the impartial spectator was merely a vehicle designed to lead to certain principles and it is these principles, not the vehicle, which define classical utilitarianism.)

Perhaps a good way to summarize Rawls' view of classical utilitarianism is by quoting a few passages from his book [19]. When discussing classical utilitarianism, Rawls says

> Social justice is the principle of rational prudence applied to an aggregative conception of the welfare of the group. (p. 27)
>
> On this conception of society, separate individuals are thought of as many different lines along which rights and duties are to be assigned and scarce means of satisfaction allocated in accordance with rules so as to give the greatest fulfillment of wants. (p. 27)
>
> [In classical utilitarianism, we think of society] as the efficient administration of social resources to maximize the satisfaction of the system of desires constructed by the impartial spectator from the many individual systems of desires accepted as given. (p. 33)

Rawls criticizes classical utilitarianism repeatedly for what he calls "the conflation of all desires into one system of desires" which he sees as "failing to take seriously the distinction between persons" [19, e.g., pp. 187–188]. Now, as Arrow [2] has pointed out, any attempt to design a coherent system of social values that would take account of the desires of all individuals is bound to contain an element of "conflation" of these desires. Nevertheless, I think that Rawls is correct to point out the failure of classical utilitarianism in coming to grips in a serious way with the phenomenon of human diversity. Human beings differ from one another in the same way that engines differ from one another, in the efficiency with which they generate energy from fuel (Edgeworth [7, p. 2]). And just as there is nothing fundamental about an engine having some specific efficiency rating, so there is nothing fundamental about a person having some specific individual characteristic. In both cases, differences among individuals matter only insofar as they can be made to work in the service of some higher principle that tells you how to allocate your scarce resources among their various potential uses. No wonder, therefore, that Rawls [19, pp. 179–180] should regard this point of view as being in conflict with Kant's imperative, that men shall treat one another not as means but as ends in themselves.

VII. Rawls' Difference Principle

Having thus rejected classical utilitarianism, Rawls sets out to propose a system of ethics that would not only "take seriously the distinction between persons" but would in fact take the distinction between persons as primary and fundamental. ("If we assume that . . . the plurality of distinct persons with separate systems of ends is an essential feature of human societies, we should not expect the principles of social choice to be utilitarian" [19, p. 29].) Moreover, Rawls' position leads hims to reject *all* forms of utilitarianism, and not just the classical variety. ("The merit of te classical view is that it clearly recognizes what is at stake, namely the relative priority of the principles of justice and of the rights derived from these principles" [19, p. 33]. This is why Rawls chooses to discuss utilitarianism mainly in its classical version. But he makes it clear that his critique is meant to apply, perhaps in varying degrees, to all forms of utilitarianism.) Rawls rejects utilitarisnism because he views the very principle of weighing one person's utility againt another's, even when it is proposed that relative weights be derived from ethical considerations, as unacceptable [19, e.g., pp. 33, 180]. For the utilitarian, whether he be classical or preclassical or modern, the concept of a *solution* for a social choice problem emerges from considerations of *trade-offs* between utilities of different persons. Thus, what Rawls sets out to do is to change the very notion of what constitutes a solution for a given problem of social choice.

Now, the *data* of any given social choice problem cannot be affected by whatever debate there might be as to what constitutes a solution of the problem. For the types of problems discussed in this essay, the data consist of a choice space, denoted X, whose elements are distribution schemes, a subset F of X that tells which distribution schemes are feasible, and, finally, a list of individuals' utility functions, $u_1, u_2,..., u_n$. Rawls' new solution concept must therefore be definable in terms of these entities. And since Rawls views "the plurality of distinct persons with separate systems of ends" as essential, it is natural that he should start by examining the utilities $u_1, u_2,..., u_n$. A sufficient condition for there to be *no* plurality of distinct persons with separate systems of ends is that all the utilities $u_1,..., u_n$ coincide on X, i.e., that for all $x \in X$ and for all pairs of individuals, i and j, the condition $u_i(x) = u_j(x)$ should hold. This means that a *necessary* condition for the existence of the plurality upon which Rawls intends to build is that there should exist a distribution scheme x and a pair of individuals, i and j, such that $u_i(x) \neq u_j(x)$. In most contexts, this condition cannot be regarded as *sufficient*, and from the inequality $u_i(x) \neq u_j(x)$ one cannot usually infer that individuals i and j are "different" in any essential way. It is therefore a shortcoming the Rawls' system that he does regard this inequality, $u_i(x) \neq u_j(x)$, as sufficient for conveying certain information on the difference between i and j. Specifically, Rawls assumes that the utilities $u_1,..., u_n$ are given in a manner that permits the inequality $u_i(x) > u_j(x)$ to be read as individual i being better off than individual j, if x is the chosen distribution scheme. (If I understand it, Rawls' discussion of "primary social goods" is meant to provide a rationale for this supposition.) This issue has been discussed extensively in the literature (e.g., Arrow [3], d'Aspremont and Gevers [6], Sen [23], and references therein). Right now, my intention is not so much to criticize Rawls' assumptions, but rather to examine his theory on its own grounds. I shall therefore accept the premise that the utilities $u_1,\ldots,u_n$ can be used to rank the relative well-being of different individuals. In Section X, I shall return to this matter and attempt to present a version of Rawls' theory where this premise is absent.

The possibility of performing an ordinal comparison between the relative well-being of different individuals means that one can speak of a widening or a narrowing of the *gap* between two individuals. Let x and y be two distribution schemes, and let u_i and u_j be the utilities of two arbitrarily chosen members of society. If it happens that the following conditions hold

$$u_i(x) > u_i(y) > u_j(y) > u_j(x), \qquad (*)$$

then one can surely say that the gap between i and j is *narrower* under y than under x. (Notice that the inequalities in (*) are written so that i is better off than j both under x and under y, but the gap is smaller under y than under x.) Suppose for a moment that society consists of these two individuals, i and j, alone and that the inequalities (*) hold. Should society pick x over y, or should it pick y over x? A utilitarian might possibly say that the answer to this question depends on the degree to which the choice between x and y affects the two individuals. If j is only slightly better off under y than under x while i is substantially better off under x than under y, then picking x over y may be justified, even though such a choice results in a wider social gap. Indeed, the utilitarian would be able to make a pretty strong case for this answer, based on continuity considerations. Such an argument might go something like this: Suppose that the inequalities (*) hold in a whole sequence of situations. Suppose also that, as we proceed along this sequence, individual j changes in such a way that the inequality $u_j(y) > u_j(x)$ tends to equality. Then, in the limit, we have $u_j(y) = u_j(x)$ and, by a strong version of the Pareto Principle, society should pick x over y. Therefore, by continuity, if the difference $u_j(y) - u_j(x)$ is sufficiently small, but positive, society should still pick x over y. When all is said and done, Rawls' system rests precisely on denying that calculations of this kind may be justified. Accordingly, he enunciates the following value judgement:

LEAST DIFFERENCE PRINCIPLE. *If society consists of individuals i and j alone, and if (*) holds, then society should never pick x over y.*

This is as weak a version as I can find for Rawls' own Difference Principle. The word "least" in Least Difference Principle can be read in two different ways, both of them appropriate. If we read "least" as modifying "difference," we get the notion that smaller differences (gaps) are better; and if we read "least" as modifying "principle," we get the notion that this is, so to speak, the least of all difference principles. (The Least Difference Principle is quite similar to Hammond's [10] Equity Axiom, but it is weaker.)

Rawls' discussion indicates that he is, in fact, prepared to make a value judgement that seems to be much stronger. Suppose that society consists of $N+1$ individuals, of whom N are just like individual i (i.e., their utilities coincide with u_i) and the last one being poor Mr. j. The integer N may be as large as we please. Suppose the inequalities (*) hold. We now have a situation where a great many individuals would benefit if x were to be chosen and only one individual—Mr. j—would benefit if y were to be chosen. Rawls' interpretation of the Kantian imperative (see Rawls [19, pp. 179–180]) leads to the value judgement that no matter how large N is, if (*) holds then society should not pick x over y. For Rawls, the denial of this judgement would constitute and instance of a person (Mr. j in this case) being treated as means and not as an end in itself. (It is not made clear, in Rawls' book, why picking x treats j as means while picking y cannot be regarded as treating i—or all those having i's utility—as means. See Harsanyi [12].)

It is easy to produce hypothetical examples where this extremely uncompromising value judgement leads to ludicrous implications (Harsanyi [12]). However, if one is willing to accept that social choice is to obey certain structural requirements (see, e.g., Arrow [3]) then it turns out that the denial of this uncompromising value judgement has consequences that are no less extreme and equally easy to expose as ludicrous. In patticular, under certain widely acceptable structural requirements, the denial of the above value judgement automatically affirms the extreme opposite judgement, namely, that even if society consists of only one individual having utility u_i and N individuals having utility u_j (reversing the situation of the previous example) and if the inequalities (*) still hold, then x should always be picked over y, no matter how large the number N.

The "structural requirements" referred to in the foregoing paragraph have been discussed extensively in the literature (Arrow [3], d'Aspremont and Gevers [6], Hammond [10], and Strasnick [26]) and there seems to be little point in dwelling upon them here once more. I shall therefore be content with the following brief and somewhat informal summary.

Let a process of social choice be called *nicely behaved* if the following five conditions are satisfied.

(1) Social choice from a finite set of feasible alternatives is always possible.

(2) Let F and F' be two feasible sets, describing two social choice situations. Assume that $F' \subset F$. If any of the schemes chosen under F are still feasible under F', then these schemes, and only they, are chosen under F'. (In particular, if under F society happens to select a unique scheme, x^*, and if $x^* \in F'$, then x^* must also be the unique selection under F'.)

(3) Consider a society with n individuals having utilities $u_1, \ldots, u_n$ and let x and y be two distribution schemes. Now let the individuals' utilities be

changed to $v_1, \ldots, v_n$ in such a way that, for all i and j, $v_i(x) \geqslant v_j(y)$ if and only if $u_i(x) \geqslant u_j(y)$. (In other words, the information concerning the well-being of individuals, when only the schemes x and y are considered, does not change, even though utilities have been changed.) Then, x is chosen over y *before* the change from $u_1, \ldots, u_n$ to $v_1, \ldots, v_n$ if and only if x is chosen over y *after* the change.

(*Note*: In the literature, conditions (2) and (3) have both come to be known as *independence of irrelevant alternatives*, which has been causing a great deal of confusion.)

(4) Let the situation be as in (3) above, but now assume that the list of utilities $v_1, \ldots, v_n$ is merely a re-arrangement (permutation) of the original list $u_1, \ldots, u_n$. If x is chosen over y *before* the change, then x is chosen over y *after* the change.

(5) Consider a society where the individuals' utilities are given by $u_1, \ldots, u_n$ and let x and y be two distribution schemes. If $u_n(x) = u_n(y)$—individual n being indifferent between x and y—then social choice between x and y must coincide with the choice between x and y is a society with $n - 1$ individuals having utilities $u_1, \ldots, u_{n-1}$.

Conditions (1) and (2) are designed to guarantee that social choice is always guided by a criterion of picking that feasible alternative which is highest in some social rank-ordering. (See Arrow [1, Theorem 3].) Condition (3) makes the social ranking of x against y dependent on information concerning x and y alone. Condition (4) guarantees equal treatment of equals, by making the information as to who has which utility irrelevant. Condition (5) requires that individuals who are indifferent between x and y not be allowed to affect the choice that society would make when forced to select between x and y.

The following proposition is a consequence of a theorem due to Hammond [10] and Strasnick [26].

PROPOSITION. *Consider a society with n individuals having utilities $u_1, \ldots, u_n$ defined over the choice space X. Let the set of feasible distribution schemes be denoted F, $F \subset X$, and let x^* be a scheme chosen from F. If social choice is nicely behaved, then, under the Pareto Principle and the Least Difference Principle, the scheme x^* must be one of the (possibly many) schemes for which the quantity*

$$\min[u_1(x), u_2(x), \ldots, u_n(x)]$$

is a maximum, for $x \in F$.

This is the so-called Maximin Rule, and the proposition can be viewed as a formalization of the process by which Rawls attempts to derive this rule from basic ethical and structural principles. Rawls argues at length that individuals gathered in an original position to select a mechanism of social choice, behind a veil of ignorance, would all agree that the mechanism to be selected should comply with the Pareto Principle and with the (Least) Difference Principle. Presumably, the parties in the original position would also agree that social choice should be nicely behaved, although this aspect is not discussed in Rawls' book. Therefore, it follows that the parties in the original position would agree, unanimously, that social choice be governed by the Maximin Rule.

VIII. ON THE EQUIVALENCE OF RAWLS AND CLASSICAL UTILITARIANISM

Consider a society with n individuals, having utilities $u_1, \ldots, u_n$. These utilities, it will be recalled, are real functions defined on some choice space, X. As before, let F be the feasible set, i.e., that subset of X from which society must actually choose, and let U be the set of feasible utility profiles, $U = \{(u_1(x), \ldots, u_n(x)) \mid x \in F\}$. The Principle of Non-polarization asserts that U is a convex set. Recall also that the *utilitarian form*, W, has been defined by writing

$$W(x, \boldsymbol{\lambda}) = \lambda_1 u_1(x) + \lambda_2 u_2(x) + \cdots + \lambda_n u_n(x)$$

for all x in X and for all $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ in the unit simplex, S.

For the sake of brevity, I shall say that a distribution scheme x^* in F *is a maxmin* if the quantity

$$\min\{u_1(x), u_2(x), \ldots, u_n(x)\}$$

has a maximum on F at $x = x^*$.

The following is a kind of duality theorem which is very close in form to the Kuhn–Tucker Theorem of non-linear programming (see, for example, Uzawa [27]). It is also obviously related to the minimax theorem for two-person zero-sum games.

THEOREM. *Let the Principle of Non-polarization be satisfied. Then, a distribution scheme x^* in F is a maxmin if, and only if, there exists a system of weights $\boldsymbol{\lambda}^* = (\lambda_1^*, \ldots, \lambda_n^*)$ in S such that the inequalities*

$$W(x, \boldsymbol{\lambda}^*) \leqslant W(x^*, \boldsymbol{\lambda}^*) \leqslant W(x^*, \boldsymbol{\lambda}) \tag{1}$$

hold for all $x \in F$ and for all $\boldsymbol{\lambda} \in S$.

When the inequalities (1) are satisfied, it is common to say that the pair

(x^*, λ^*) is a *saddlepoint* of the utilitarian form $W(x, \lambda)$, for $x \in F$ and $\lambda \in S$. Using this terminology, what the theorem says is that, under non-polarization, x^* is a maxmin if and only if x^*, together with some system of weights λ^*, is a saddlepoint of the utilitarian form, W.

The proof of the foregoing theorem relies on the separation theorem for convex sets, and it is virtually the same as the proof of the minimax theorem for two-person zero-sum games. So, there is no point in reproducing the proof here.

Somewhere on page 12 of Edgeworth's book one would have expected to find a sentence that would read roughly as follows: "Just as mechanical action is governed by the law that the Hamiltonian Form shall be at a saddlepoint, so social action is governed by the law that the Utilitarian Form shall be at a saddlepoint." Edgeworth clearly had something like this in mind, so if he had actually said it, my task of showing that (under non-polarization) Rawls' theory is equivalent to Edgeworth's would now be complete. But Edgeworth did not quite say it, so completing my task is going to require some writing.

One thing ought to be clear: If the pair (x^*, λ^*) is a saddlepoint of W, that is, if the inequalities

$$W(x, \lambda^*) \leqslant W(x^*, \lambda^*) \leqslant W(x^*, \lambda)$$

hold for all $x \in F$ and $\lambda \in S$, then the pair (x^*, λ^*) is a solution of the problem

$$\min_{\lambda \in S} \max_{x \in F} W(x, \lambda).$$

Therefore, the foregoing theorem tells us right away that a distribution scheme, say x^*, selected from the feasible set F by Rawls' Difference Principle must, under non-polarization, satisfy the classical utilitarians' rule for social choice (see Section VI). In other words, under non-polarization, if a Rawlsian were to recommend some social action, based on Rawls' Difference Principle, then a classical utilitarian would always concur.

Now, what about the converse? Would a Rawlsian always be inclined to concur in a recommendation made by a classical utilitarian? The answer is yes, *provided* one takes into account the classical utilitarians' rule for breaking ties. Consider a problem of social choice for which a solution, according to Rawls' Maxmin Rule, exists (so the question posed above is meaningful). By the theorem, this Rawlsian solution would have to be *one* of the solutions that a classical utilitarian would recommend. In particular, if the social choice rule of classical utilitarianism leads to a unique solution, then this solution must coincide with the Rawlsian solution, i.e., the Rawlsian would concur with the utilitarian. However, if the choice rule of classical utilitarianism leads to *several solutions, locked in a tie, then all we know is that one of these would be Rawlsian, but not necessarily all. Thus, it appears as though, in case of a tie, a classical utilitarian might end up making a recommendation in which a Rawlsian would not concur.*

EXAMPLE. *Take a two-person case, and let the set U of feasible utility* profiles consist of all pairs (u_1, u_2) of non-negative utility numbers satisfying $u_1 + u_2 \leqslant 1$. *In this case, classical utilitarianism leads to the result that all* the Pareto-efficient utility profiles—i.e., all feasible pairs (u_1, u_2) satisfying $u_1 + u_2 = 1$—are locked in a tie. Among these, the only configuration satisfying Rawls' Maxmin Rule is the one for which $u_1 = u_2 = \frac{1}{2}$.

But classical utilitarians were fully aware that, in cases where the principle of greatest happiness results in a rie, a further principle is needed to decide the issue. Here is the appropriate passage from Sidgwick [24, 3rd ed., pp. 413–414] on this point:

> It is evident that there may be different ways of distributing the same quantum of happiness among the same number of persons; in order, therefore, that the Utilitarian criterion of right conduct may be as complete as possible, we ought to know which of these ways is to be preferred ... Now the Utilitarian formula seems to supply no answer to this question: at least we have to supplement the principle of seeking the reatest happiness on the whole by some principle of Just or Right distribution of this happiness. The principle which most Utilitarians have either tacitly or expressly adopted is that of pure equality.

Thus, classical utilitarianism amounts to the use of two principles, arranged in a hierarchy: The principle of greatest happiness should be applied first. If this principle is not sufficient, in the sense that it leads to several solutions locked in a tie, then the principle of greatest equality should apply, to break the tie. (In the simple example given above, the principle of greatest happiness led to all non-negative pairs (u_1, u_2) satisfying $u_1 + u_2 = 1$ as being equally good. Therefore, by the principle of greatest equality, the tie-breaking pair is given by $u_1 = u_2 = \frac{1}{2}$, which coincides with the solution according to the Maxmin Rule.) A moment's reflection reveals that, in the presence of the theorem stated above, this hierarchy of principles always leads, under non-polarization, to a selection that conforms with Rawls' Maxmin Rule.

IX. ON DROPPING NON-POLARIZATION

The equivalence of Rawls' theory of distribution and that of classical utilitarianism rests heavily on the assumption of non-polarization. This assumption is justified when it is believed either that society can make randomized choices or else that goods are smoothly transferable among individuals, that social policies are capable of gradual implementation, and that individuals are subject to some form of diminishing marginal utility. Conceivably, under some circumstances, these beliefs may be unwarranted. When this is the case, i.e., when it is not reasonable to suppose tht the set U of feasible utility profiles is convex, the difference between Rawl' and classical tuilitarianism springs up and becomes pronounced. Without nonpolarization, classical utilitarianism tends to push social choice towards extremes, whereas Rawls' Difference Principle tends to work in favor of moderation, to the extent that moderation is feasible. A simple example may serve to illustrate this contrast.

Suppose that society consists of two individuals, having utilities u_1 and u_2, and assume that the set U of feasible utility profiles is as depicted in Fig. 1. In this situation, Rawls' Maxmin Rule would select the utility profile B. On the other hand, classical utilitarianism, guided by the rule $\min_\lambda \max_x W(x, \lambda)$, would select the profiles A and C indifferently, without being able to decide between them. (Applying to the principle of greatest equality, as a tie-breaker, is clearly of no use in a comparison between A and C.)

Another case where non-polarization does not hold is one where the number of options open to society is finite. In practice, of course, society never faces a choice among infinitely many alternatives, and the question of finitely versus infinitely many alternatives is one of selecting a *model.* If the situation is such that the finie-options model is more appropriate, then non-polarization fails so that Rawls' Maxmin and classical utilitarianism diverge. Moreover, it can be shown that Rawls' theory is, in this case, more powerful than classical utilitarianism. By "more powerful" I mean that Rawls' theory results in an unambiguous choice more often than does classical utilitarianism. Once again, a simple example, with two individuals having utilities u_1 and u_2, may serve to illustrate the point. Suppose that society faces a choice between two isolated policies leading, respectively, to the

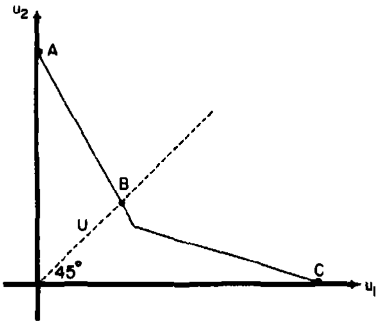


FIGURE 1

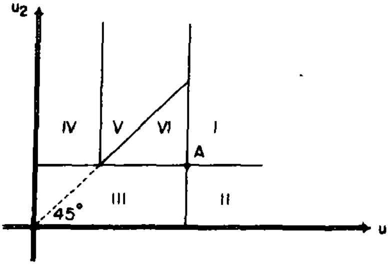


FIGURE 2

utility profiles A and B. In Fig. 2, the utility profile A has been marked and then, using the point A, the non-negative quadrant has been divided into six regions marked by Roman numerals, I through VI. If the utility profile B falls in one of the regions marked I, II, III, or VI, then both Rawls' Maxmin and classical utilitarianism are capable of deciding the issue and, in fact, the choices made by the two theories are the same (in this simple two-alternative case). However, if B happens to fall either inside IV or inside V, then classical utilitarianism is not capable of deciding between A and B, while Rawls' Maxmin is. Specifically, Rawls' Maxmin will pick A over B if B lies in IV and will pick B over A if B lies in V. The only case where neither theory is capable of deciding the issue is the one where B happens to fall right on the vertical line separating IV and V. (If B happens to fall on the horizontal line passing through A, then both Rawls' and classical utilitarianism require the *strong* version of the Pareto Principle in order to decide between A and B.) In this sense, it may be said that Rawls' Maxmin is more powerful than classical utilitarianism. Now, when the number of isolated alternatives is greater than 2, the situation becomes much more complicated. However, it is still true that in all configurations where classical utilitarianism leads to a unique choice so does Rawls' Maxmin, but not conversely.

X. RAWLS' DIFFERENCE PRINCIPLE WITHOUT INTERPERSONAL COMPARISONS

Before proceeding, let me reinstate the Principle of Non-polarization, which will be assumed to hold throughout the remainder of this essay: The set U of feasible utility profiles is convex.

As has been pointed out in Section VII above, the theory of distribution proposed by Rawls [19] rests on the stipulation that assertions of the form $u_i(x) > u_j(y)$ are meaningful. (Here, as always, $u_1, \ldots, u_n$ are the utilities of the individuals who make up the society under consideration.) Specifically, the inequality $u_i(x) > u_j(y)$ is shorthand for the statement "individual i is better off under the distribution scheme x than individual j under the distribution scheme y." Now, this element of Rawls' theory is regarded by most people as a significant shortcoming because the meaningfulness of an inequality of the form $u_i(x) > u_j(y)$ implies that u_i and u_j—and therefore all the utilities $u_1,..., u_n$—are measured in the same units. Not only is your utility assumed to be the same entity as my utility, the units used to measure the two utilities are assumed to be the same. These common units can be used to construct a procedure for "objective" measurement of well-being, and this is a conclusion to which utility theory (whether in ordinalist or in cardinalist form) simply cannot lend any support.

In this section, I propose to reformulate Rawls' Difference Principle in a manner that does not involve the supposition that utilities are interpersonally comparable. The tenets of utility theory—albeit in its cardinalist version—will be adhered to fully. To achieve this minor feat, I shall appeal to the conceptual framework, due to Shapley [25], that has been outlined in Section V.

The basic idea is quite simple, and it can be summarized in a few sentences. With every Pareto-efficient distribution scheme, one can associate a system of *implicit* conversion rates that transform individual utilities into some common unit of measurement. This common unit of measurement can be used to perform interpersonal comparisons even though the individuals' utilities themselves are not comparable. In Shapley's [25] terminology, one could say this: Let us pretend for a moment that utilities *are* transferable among individuals and that the rates of transfer are given by the weights $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$. That is, we pretend that when person i gives up a unit of his utility, u_i, to be transferred to person j, the gain to person j is λ_i/λ_j units of his own utility, u_j. (Here, all the weights λ_i, for $i = 1,..., n$, are assumed to be positive.) In this kind of never-never land (to use Shapley's phrase) Rawls' Difference Principle would call for a distribution of utility levels $u_1,..., u_n$ that would maximize the least of the quantities $\lambda_1 u_1, \lambda_2 u_2,..., \lambda_n u_n$. But now, seeing that utility is *not* in fact transferable, what we ought to do, according to Shapley, is to keep changing the weights $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$ until we hit upon a particular system of weights, say $\boldsymbol{\lambda}^* = (\lambda_1^*,..., \lambda_n^*)$, having the property that the distribution $(u_1^*,..., u_n^*)$ of utility levels that maximizes the least of the numbers $\lambda_1^* u_1^*,..., \lambda_n^* u_n^*$ (in the transferable utility game defined by $\boldsymbol{\lambda}^*$) can be achieved *without* utility transfers actually being necessary.

To get started, I shall assume, in keeping with Rawls' contractarian framework, that no distribution scheme can be put into effect unless *all* individuals agree. Every individual has the power to veto a proposed distribution scheme, thereby forcing society to stay at the status quo. Staying at the status quo (or, if one prefers, at the "state of nature") is always a feasible option, which will be denoted x_0. In the absence of interpersonal comparability, one may assume that the individuals' utilities have been normalized so as to satisfy $u_1(x_0) = u_2(x_0) = \cdots = u_n(x_0) = 0$. Under this normalization, a distribution scheme x is relevant for social choice only if the utility levels $u_1(x),..., u_n(x)$ are all non-negative. Therefore, I will assume henceforth that the set F of feasible distribution schemes has the property that $x \in F$ implies $u_i(x) \geqslant 0$ for $i = 1,..., n$. Furthermore, I shall assume that there exists at least one social contract where all individuals are actually better off than they would be if society were to stay at the status quo. In symbols, this assumption asserts that there exists a feasible distribution scheme $\hat{x}$ ($\hat{x} \in F$) such that $u_i(\hat{x}) > 0$ for $i = 1,..., n$. Obviously, if this assumption fails, then the Rawlsian notion of maximizing the well-being of the least advantaged individual degenerates into a triviality, because in this case staying at the status quo is always a solution.

Consider the problem of choosing a distribution scheme from a feasible set F, and let x^* belong to F. If $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$ is an n-tuple of non-negative weights satisfying $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$ and if the utilitarian form $W(x, \boldsymbol{\lambda})$ is a maximum on F when $x = x^*$, then I shall say that $\boldsymbol{\lambda}$ is a system of weights *corresponding* to the scheme x^*. Let $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$ be a system of weights corresponding to some (Pareto-efficient) distribution scheme x^*, and let x be any distribution scheme. Clearly, the magnitudes $\lambda_1 u_1(x),..., \lambda_n u_n(x)$ are all measured in the same unit. Moreover, if $\boldsymbol{\mu} = (\mu_1,..., \mu_n)$ is a system of weights corresponding to some other scheme, say y^*, then the magnitudes $\mu_1 u_1(x),..., \mu_n u_n(x)$ are still measured in the same unit as were the earlier magnitudes, $\lambda_1 u_1(x),..., \lambda_n u_n(x)$. Therefore, as far as units of measurement are concerned, it is meaningful to compare a quantity of the form $\lambda_i u_i(x)$ with another quantity of the form $\mu_j u_j(y)$, where x and y are any distribution schemes. From a purely formal point of view, the inequality

$$\lambda_i u_i(x^*) > \mu_j u_j(y^*)$$

is a meaningful assertion. The question, however, is: Under the assumption that $(\lambda_1,..., \lambda_n)$ and $(\mu_1,..., \mu_n)$ are systems of weights corresponding, respectively, to x^* and y^*, can this inequality be taken to mean that individual i is better off under x^* than individual j under y^*? In order to answer this question, one must start by examining the case of *intra*personal comparisons, where we take $i = j$. That is, if $\boldsymbol{\lambda} = (\lambda_1,..., \lambda_n)$ and $\boldsymbol{\mu} = (\mu_1,..., \mu_n)$ are systems of weights corresponding, respectively, to x^* and y^*, and if i is some integer satisfying $1 \leqslant i \leqslant n$, can one infer from the inequality

$$\lambda_i u_i(x^*) > \mu_i u_i(y^*)$$

that individual i *prefers* x^* over y^*? Is there any necessary connection between the statement $\lambda_i u_i(x^*) > \mu_i u_i(y^*)$ and the statement $u_i(x^*) > u_i(y^*)$? In the special case of a society that consists of two individuals, the answer to this question is simple and straightforward: If $n = 2$, then the inequality

$$(\lambda_i u_i(x^*) - \mu_i u_i(y^*))(u_i(x^*) - u_i(y^*)) \geqslant 0$$

is true for $i = 1, 2$, with equality holding if and only if either $u_i(x^*) = u_i(y^*)$ or $\lambda_i = \mu_i = 0$. In other words, when $n = 2$, the two differences—$\lambda_i u_i(x^*) - \mu_i u_i(y^*)$ and $u_i(x^*) - u_i(y^*)$—cannot have opposite signs, so it is impossible to have $\lambda_i u_i(x^*) > \mu_i u_i(y^*)$ with the ith individual's preference between x^* and y^* running the other way. (This assertion is implied by the more general assertion that will be stated and proved a few lines below.)

If there are more than two individuals, then the quantities $\lambda_i u_i(x^*) - \mu_i u_i(y^*)$ and $u_i(x^*) - u_i(y^*)$ may have opposite signs, but this has to do with the relationships among individuals other than i himself. If one controls for these relationships, then the desired result emerges again. One way to control for the relationships among individuals other than i is to suppose that $(\lambda_j/\lambda_k) = (\mu_j/\mu_k)$ for all $j \neq i$ and $k \neq i$. What this means is that in the transition from x^* to y^*, rates of conversion between utilities of individuals other than the ith individual are unaffected. Another way, which also leads to the desired result, is to suppose that $(u_j(x^*)/u_k(x^*)) = (u_j(y^*)/u_k(y^*))$ for all $j, k \neq i$, which means that relative welfare positions of individuals other than i do not change in the transition from x^* to y^*. Here, I shall be content to spell out only the first of these two alternatives. The following assertion is true: Let i be an integer satisfying $1 \leqslant i \leqslant n$ and let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ be systems of weights corresponding, respectively, to x^* and y^*. If the equation

$$\left(\lambda_j \middle/ \sum_{k \neq i} \lambda_k\right) = \left(\mu_j \middle/ \sum_{k \neq i} \mu_k\right) \tag{2}$$

holds true for all $j \neq i$, or if one of the denominators in Eq. (2) vanishes, then the inequality

$$(\lambda_i u_i(x^*) - \mu_i u_i(y^*))(u_i(x^*) - u_i(y^*)) \geqslant 0. \tag{3}$$

holds, with equality if, and only if, either $u_i(x^*) = u_i(y^*)$ or $\lambda_i = \mu_i = 0$. Note that, in the case $n = 2$, the assumption on the relative weights, (2), is always satisfied (trivially) so the previous assertion is entailed in this one.

Proof of the assertion. From the fact that $\boldsymbol{\lambda}$ corresponds to x^* and $\boldsymbol{\mu}$ corresponds to y^*, it follows that two inequalities may be written:

$$\sum_{j=1}^{n} \lambda_j u_j(x^*) \geqslant \sum_{j=1}^{n} \lambda_j u_j(y^*), \tag{4}$$

$$\sum_{j=1}^{n} \mu_j u_j(x^*) \leqslant \sum_{j=1}^{n} \mu_j u_j(y^*). \tag{5}$$

If $\lambda_i = 1$, then (3) follows from (4) and if $\mu_i = 1$, then (3) follows from (5). This takes care of the case where one of the denominators in (2) vanishes. Now assume that $\lambda_i < 1$ and $\mu_i < 1$ and divide (4) and (5), respectively, by $1 - \lambda_i$ and $1 - \mu_i$. This will lead to

$$\sum_{j=1}^{n} \left(\frac{\lambda_j}{1-\lambda_i} - \frac{\mu_j}{1-\mu_i}\right)(u_j(x^*) - u_j(y^*)) \geqslant 0. \tag{6}$$

By (2), all the terms appearing in (6) vanish, except for the ith term. Thus, one gets

$$\left(\frac{\lambda_i}{1-\lambda_i} - \frac{\mu_i}{1-\mu_i}\right)(u_i(x^*) - u_i(y^*)) \geqslant 0, \tag{7}$$

which holds if, and only if

$$(\lambda_i - \mu_i)(u_i(x^*) - u_i(y^*)) \geqslant 0. \tag{8}$$

And since all the quantities appearing in (8) are non-negative—recall that $u_i(x) \geqslant 0$ for all $x \in F$—the desired result, (3), follows readily. ∎

We see, then, that under a certain *ceteris paribus* assumption, the inequality $\lambda_i u_i(x^*) > \mu_i u_i(y^*)$ does indeed admit an interpretation showing a preference for x^* over y^* by individual i. Now, what about the inequality $\lambda_i u_i(x^*) > \mu_j u_j(y^*)$, where $i \neq j$? Here, the only question that one can ask is the following. Suppose, for the sake of argument, that u_i and u_j have somehow been pre-calibrated, for some $i \neq j$, in such a way that the inequality $u_i(x^*) > u_j(y^*)$ is meaningful. *In this case*, will it also be true that $\lambda_i u_i(x^*) > \mu_j u_j(y^*)$? Clearly, if the inequality $u_i(x^*) > u_j(y^*)$ is meaningful, then u_i and u_j must be measured in the same unit which means, in turn, that the rate at which u_i can be converted into any universal unit must coincide with the rate at which u_j can be converted into that unit. Therefore, $\lambda_i = \mu_j$ and the inequality $\lambda_i u_i(x^*) \geqslant \mu_j u_j(y^*)$ must hold, with a strict inequality in all but the extreme case, when $\lambda_i = \mu_j = 0$.

These considerations lead me to conclude that it is not unreasonable to try and reformulate Rawls' Difference Principle on the basis of comparisons of the type $\lambda_i u_i(x^*) > \mu_j u_j(y^*)$, where $(\lambda_1, \ldots, \lambda_n)$ and $(\mu_1, \ldots, \mu_n)$ are systems of weights that correspond, respectively, to x^* and y^*. This can be done formally, using Shapley's procedure, which has two stages. The first stage consists of specifying a rule for distribution in the case of transferable utility. The second stage consists of allowing the rates of transfer of utility to vary until the distribution problem is such that its solution can be achieved *without* utility transfers actually taking place. These two stages will now be taken up in order and described more fully.

In the first stage, a hypothetical set of feasible utility profiles, to be denoted $U(\boldsymbol{\lambda}, K)$, is defined by writing $(v_1,\dots, v_n) \in U(\boldsymbol{\lambda}, K)$ if and only if $v_1,\dots, v_n$ are non-negative real numbers satisfying $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \leqslant K$. Here, $\boldsymbol{\lambda} = (\lambda_1,\dots, \lambda_n)$ is any system of weights such that $\lambda_i > 0$ for $i = 1,\dots, n$ and $\lambda_1 + \cdots + \lambda_n = 1$, and K is some non-negative real number. The weights $\lambda_1,\dots, \lambda_n$ must all be strictly positive because $\lambda_i = 0$ for some i would mean that utility transfers to and from the ith individual are not possible. The constant K represents the maximum amount of welfare that can be achieved by society as a whole in this hypothetical situation, and it is postulated that "welfare" is an entity which can be divided up among individuals in any way, with the weights $\lambda_1,\dots, \lambda_n$ acting as the rates of conversion of individual utilities into units of this common entity, called "welfare." Now let $(v_1,\dots, v_n)$ be a feasible utility profile, $(v_1,\dots, v_n) \in U(\boldsymbol{\lambda}, K)$. Under this profile, person i's well-being, measured in common units of welfare, is given by the quantity $\lambda_i v_i$. Therefore, in this hypothetical setting, Rawls' Difference Principle would dictate that $(v_1,\dots, v_n)$ be selected from the set $U(\boldsymbol{\lambda}, K)$ so as to maximize the *least* of the quantities $\lambda_1 v_1,\dots, \lambda_n v_n$. If $(v_1^*,\dots, v_n^*)$ is a utility profile in $U(\boldsymbol{\lambda}, K)$ at which this maximum is attained, then I shall say, for the sake of brevity, that $(v_1^*,\dots, v_n^*)$ is a *λ-maxmin utility profile*, for total welfare level K.

In the second stage of the argument, transferable utility is no longer postulated, which means that the set of feasible utility profiles is no longer of the special form, $U(\boldsymbol{\lambda}, K)$, of the previous paragraph. Instead, we are back with our old set U, which is obtained by taking the image of the set F of feasible distribution schemes in the space of utility profiles, $U = \{(u_1(x),\dots, u_n(x)) \mid x \in F\}$. To handle this situation, Shapley tells us to keep searching through different systems of weights, $\boldsymbol{\lambda}$, locating in each case a $\boldsymbol{\lambda}$-maxmin utility profile for some appropriate welfare level K. The search is over as soon as one finds a system of weights, say $\boldsymbol{\lambda}^*$, having the following two properties: First, there must exist a feasible distribution scheme x^* ($x^* \in F$) such that the n-tuple $(u_1(x^*),\dots, u_n(x^*))$ is a $\boldsymbol{\lambda}$-maxmin utility profile, with the constant K to be determined by the next requirement. This means that at $\boldsymbol{\lambda}^*$, the transferable utility maxmin is attainable without any transfers actually taking place. Second, the system of weights $\boldsymbol{\lambda}^*$ must correspond to the distribution scheme x^* so that, at x^*, utility is in fact *locally* transferable at the rates implied by $\boldsymbol{\lambda}^*$. (More precisely, $W(x, \boldsymbol{\lambda}^*)$ must be a maximum on F at $x = x^*$, which means that if person i's utility is reduced from $u_i(x^*)$ to $u_i(x^*) - \varepsilon$, then the maximum F-feasible utility gain that person j can get from this is no greater than $\varepsilon(\lambda_i^*/\lambda_j^*)$.) This second requirement determines the value of the constant K appearing in the first requirement. Specifically, $K = \sum \lambda_i^* u_i(x^*) = W(x^*, \boldsymbol{\lambda}^*)$.

All this may now be summarized by a formal definition that will lead to a statement of the ethical principle which has been the object of the present discussion.

DEFINITION. Let F, $F \subset X$, be the set of feasible distribution schemes and let $u_1, \dots, u_n$ be utilities defined on X. A distribution scheme $x^* \in F$ will be referred to as a *Rawls–Shapley scheme* if there exists a system of weights $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_n^*)$, with $\lambda_i^* > 0$ for $i = 1, \dots, n$ and $\lambda_1^* + \cdots + \lambda_n^* = 1$, such that the following two conditions hold:

(i) $\boldsymbol{\lambda}^*$ is a system of weights corresponding to x^*. That is,

$$W(x^*, \boldsymbol{\lambda}^*) \geqslant W(x, \boldsymbol{\lambda}^*) \qquad \text{for all} \quad x \in F;$$

(ii) the n-tuple $(u_1(x^*), \dots, u_n(x^*))$ is a $\boldsymbol{\lambda}^*$-maxmin utility profile for total welfare level $W(x^*, \boldsymbol{\lambda}^*)$. In other words, the quantity

$$\min[\lambda_1^* v_1, \lambda_2^* v_2, \dots, \lambda_n^* v_n]$$

is at a maximum on the set $U(\boldsymbol{\lambda}^*, W(x^*, \boldsymbol{\lambda}^*))$ when $v_1 = u_1(x^*)$, $v_2 = u_2(x^*), \dots, v_n = u_n(x^*)$.

To spare the reader the need to flip pages back and forth, let me repeat the definition of the set $U(\boldsymbol{\lambda}, K)$:

$$U(\boldsymbol{\lambda}, K) = \left\{ (v_1, \dots, v_n) \;\middle|\; \begin{array}{l} v_1, \dots, v_n \text{ non-negative real} \\ \text{numbers with } \lambda_1 v_1 + \cdots + \lambda_n v_n \leqslant K \end{array} \right\}.$$

It is now possible to state:

DIFFERENCE PRINCIPLE WITHOUT INTERPERSONAL COMPARISONS. *Society should always select a Rawls–Shapley Distribution scheme.*

An analysis of this new principle would require that three questions be considered. First, under what conditions do Rawls–Shapley distribution schemes exist? Second, if Rawls–Shapley schemes exist, are they typically unique? (If they are not, then a tie-breaking rule may be needed.) Third, what is the axiomatic basis for a principle of social choice demanding that Rawls–Shapley schemes shall always be selected? The following theorem serves to answer all three questions at once.

THEOREM. *Let F $(F \subset X)$ be the set of feasible distribution schemes and* *let the Principle of Non-polarization be satisfied. Assume also that there exists in F a distribution scheme $\hat{x}$ with $u_i(\hat{x}) > 0$ for $i = 1,\dots, n$. A distribution scheme x^*, $x^* \in F$, is a Rawls–Shapley scheme if, and only if, the product*

$$u_1(x) \cdot u_2(x) \cdots u_n(x)$$

is a maximum on F when $x = x^*$.

In other words, x^* is a Rawls–Shapley scheme if, and only if, x^* is the solution on F of Nash's [17] bargaining problem. (Recall that utilities have all been so normalized that $u_i(x_0) = 0$ for $i = 1,\dots, n$, where x_0 is the so-called status quo.)

Sketch of the proof. One proceeds by showing, first of all, that x^* is a Rawls–Shapley scheme if, and only if, there exists a system of weights $\boldsymbol{\lambda}^* = (\lambda_1^*,\dots, \lambda_n^*)$ with $\lambda_i^* > 0$ for $i = 1,\dots, n$ and $\lambda_1^* + \cdots + \lambda_n^* = 1$, such that the following two conditions hold:

$$\text{(a)} \quad W(x^*, \boldsymbol{\lambda}^*) \geqslant W(x, \boldsymbol{\lambda}^*) \qquad \text{for all} \quad x \in F;$$

$$\text{(b)} \quad \lambda_1^* u_1(x^*) = \lambda_2^* u_2(x^*) = \cdots = \lambda_n^* u_n(x^*).$$

Showing this is quite straightforward. Now, shifting the discussion into the space of utility profiles, it is possible to write down the problem of finding a Rawls–Shapley scheme as follows: Let U be a subset of the non-negative orthant in Euclidean n-space, and assume that U is convex and that it possesses a point whose components are strictly positive. Find a point $\mathbf{u}^* = (u_1^*,\dots, u_n^*)$ for which there exists a system of weights $\boldsymbol{\lambda}^* = (\lambda_1^*,\dots, \lambda_n^*)$ with $\lambda_i^* > 0$ for $i = 1,\dots, n$ and $\lambda_1^* + \cdots + \lambda_n^* = 1$, such that

$$\text{(a}'\text{)} \quad \sum_{i=1}^{n} \lambda_i^* u_i^* \geqslant \sum_{i=1}^{n} \lambda_i^* u_i \qquad \text{for all} \quad (u_1,\dots, u_n) \in U;$$

$$\text{(b}'\text{)} \quad \lambda_1^* u_1^* = \lambda_2^* u_2^* = \cdots = \lambda_n^* u_n.$$

What remains now is to verify that we have here a procedure for associating a point $\mathbf{u}^*$ with a given convex set U in such a way that all the axioms of Nash's [17] solution of the bargaining problem are satisfied. (See, e.g., Roth [20] for a statement of these axioms in the case of a bargaining problem involving n persons.) This verification completes the proof of the theorem, in view of the fact that, under the hypotheses of the theorem, if a solution to Nash's problem exists, then it is unique.

Here, then, are the answers to the three questions posed above. First, under the premises of the theorem, a sufficient condition for the existence of the Rawls–Shapley scheme is the compactness of the set U of feasible utility profiles (and boundedness of U is in fact a necessary condition). Second, under the presmises of the theorem, if a Rawls–Shapley scheme exists then it is unique, up to a matter of indifference for all individuals. Third, the axiomatic basis for selecting a Rawls–Shapley scheme is given by Nash's four well-known postulates, namely, Pareto-efficiency, symmetry, invariance under linear transformations of utilities, and invariance under deletion of non-optimal alternatives.

There is an air of anti-climax about the foregoing theorem: All this work to release Rawls from the clutches of interpersonal comparisons, culminating in a return to a conceptual framework introduced by Nash in 1950. This is particularly disappointing in view of the fact that Nash's bargaining theory is known to have a serious deficiency—when interpreted as a theory of distributive justice—namely, that it leads to distributions that tend to assign greater rewards to persons having utilities that are closer to linearity, and this can be regarded as a bias in favor of the rich and against the poor. (See, for example, Luce and Raiffa [14, pp. 129–130].)

However, it is possible also to turn the matter around and state the moral of the foregoing theorem in a more positive way: We find John Nash's view of how bargaining determines a social contract receiving renewed support from unexpected quarters, namely, from John Rawls' ethics (combined with Lloyd Shapley's method for dealing with interpersonal comparisons). It is possible that a change in the definition of "the status quo" would permit this positive conclusion to stand, while at the same time eliminating the problem of the bias in favor of the rich. But this would be a topic for a separate study.

REFERENCES

1. K. J. ARROW, Rational choice function and orderings, *Economica, New Series*, **26** (1959).
2. K. J. ARROW, Some ordinalist-utilitarian notes on Rawls' theory of justice, *J. Philos.* **70**, No. 9 (1973).
3. K. J. ARROW, Extended sympathy and the possibility of social choice, *Amer. Econ. Rev.* **67**, No. 1 (1977)
4. R. J. AUMANN AND M. KURZ, Power and taxes, *Econometrica* **45**, No. 5 (1977).
5. B. BARRY, "The Liberal Theory of Justice," Oxford Univ. Press (Clarendon), London/New York, 1973.
6. C. D'ASPREMONT AND L. GEVERS, Equity and the informational basis of collective choice, *Rev. Econ. Studies* **44**, No. 2 (1977).
7. F. Y. EDGEWORTH, "Mathematical Physics," Kegan Paul, London, 1881. (Reprinted 1967 by Augustus M. Kelley, New York.)
8. F. Y. EDGEWORTH, "Papers Relating to Political Economy," Vol. III, Macmillan & Co., London, 1925.
9. S. GORDON, John Rawls' Difference Principle, utilitarianism, and the optimum degree of inequality, *J. Philos.* **70**, No. 9 (1973).
10. P. J. HAMMOND, Equity, Arrow's conditions, and Rawls' Difference Principle, *Econometrica* **44**, No. 4 (1976).
11. J. C. HARSANYI, Cardinal welfare, individualistic ethics and interpersonal comparisons of utility, *J. Political Econ.* **63** (1955).

12. J. C. HARSANYI, Can the Maximin Principle serve as a basis for morality? A critique of John Rawls' theory, *Amer. Political Sci. Rev.* **69** (1975).
13. D. H. KRANTZ, R. D. LUCE, P. SUPPES, AND A. TVERSKY, "Foundations of Measurement," Vol. 1, Academic Press, New York, 1971.
14. R. D. LUCE AND H. RAIFFA, "Games and Decisions," Wiley, New York, 1958.
15. D. LYONS, Rawls versus utilitarianism, *J. Philos.* **69**, No. 18 (1972).
16. G. MCCLELLAND AND J. ROHRBAUGH, "Who Accepts the Pareto Axiom? The Role of Utility and Equity in Arbitration Decisions," Research Report No. 202, Institute of Behavioral Science, University of Colorado, 1978.
17. J. F. NASH, The Bargaining problem, *Econometrica* **18** (1950).
18. R. NOZICK, "Anarchy, State and Utopia," Basic Books, New York, 1974.
19. J. RAWLS, "A Theory of Justice," Belknap Press, Cambridge, Mass., 1971.
20. A. E. ROTH, The Nash solution and the utility of Bargaining, *Econometrica* **46**, (1978).
21. A. K. SEN, "Collective Choice and Social Welfare," Holden–Day, San Francisco, 1970.
22. A. K. SEN, Rawls versus Bentham: An axiomatic examination of the pure distribution problem, *Theory & Decision* **4** (1974). (Reprinted *in* "Reading Rawls" (N. Daniels, Ed.), Basic Books, New York, 1974).
23. A. K. SEN, On Weights and measures: Informational constraints in social welfare analysis, *Econometrica* **45**, No. 7 (1977).
24. H. SIDGWICK, "The Methods of Ethics," Macmillan & Co., London, 1874.
25. L. S. SHAPLEY, Utility comparison and the theory of games, *in* "La Décision," Colloques Internationaux du CNRS, CNRS, Paris, 1969.
26. S. STRASNICK, Social Choice and the derivation of Rawls' Difference Principle, *J. Philos.* **73**, No. 4 (1976).
27. H. UZAWA, The Kuhn–Tucker theorem in concave programming, *in* "Studies in Linear and Non-Linear Programming" (K. J. Arrow, L. Hurwicz, and H. Uzawa, Eds.), Stanford Univ. Press, Stanford, Calif., 1958.

JOURNAL OF ECONOMIC THEORY **24**, 40–47 (1981)

Effective Nondictatorial Domains

KI HANG KIM AND FRED W. ROUSH

Mathematics Research Group, Alabama State University, Montgomery, Alabama 36101

Received August 27, 1979; revised April 30, 1980

The basic setting of this paper is similar to that of Kalai and Muller [1]. Suppose we have n voters and m alternatives. Each voter's preferences here will be a linear order on $\{1, 2,..., m\}$, the set of alternatives.

DEFINITION. Let L denote the set of all linear orders on $\{1, 2,..., m\}$. A *domain* is a subset of L. Let O_W denote the set of weak orders on $\{1, 2,..., m\}$.

DEFINITION. A *social welfare function* on a domain D is a function from D^n to O_W.

DEFINITION. A social welfare function F is *dictatorial* if there exists i such that for all $P_1,..., P_n \in D$, $F(P_1,..., P_n) = P_i$. Voter i is said to be the *dictator*.

DEFINITION. A social welfare function F is *binary* if and only if for any pair x, y of alternatives, and any $P_1,..., P_n$, $Q_1,..., Q_n \in D$ the following is true: if for all i, xP_iy if and only if xQ_iy then $xF(P_1,..., P_n)y$ if and only if $xF(Q_1,..., Q_n)y$.

DEFINITION. A social welfare function F is *Pareto* if and only if for any alternatives x, y and for any $P_1,..., P_n \in D$ if xP_iy for all i then $xF(P_1,..., P_n)y$.

In the following F will always denote a Pareto, binary, nondictatorial social welfare function, on the domain D (or some other specified domain).

A notation such as $D(m)$ will indicate a domain for m alternatives. Note that unlike Kalai and Muller, we allow the value of a socal welfare function to be any weak order, not a linear order. However, the preferences of individual voters are to be linear orders.

Kalai and Muller found necessary and sufficient conditions that a nondictatorial, Pareto, binary, single-valued social welfare function exist on a given domain. How large can the domain be? By Arrow's theorem, D must be a

0022-0531/81/010040-08$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

proper subset of L. It has long been known that the domain of single-peaked preferences does have such a social welfare function. Kalai and Ritz [2] found a class of such domains, having

$$\frac{m!}{2} + (m-1)!$$

Accession Number.
....
Date..........................

elements. These are called *domains with an inseparable pair.*

Definition. The pair (x, y) is *inseparable in the domain D* if and only if *in every ordering P of D if xPy then there does not exist z such that xPz and zPy.*

Kim and Roush [3] have shown that no nondictatorial domain has larger size than this. However domains with an inseparable pair are rather special in that, with the known social welfare function, one individual can dictate any choice except the choice between x and y. Thus he is almost a dictator. Here we study the case where no individual is a partial dictator of this general type and show that the size of the domain must be small as compared with L. This indicates that on "most domains" every social welfare function has one voter who is a partial dictator.

Definition. For a social welfare function F, voter *i controls a choice* (x, y) if and only if for all $P_1,\dots, P_n \in D$, xP_iy implies $xF(P_1,\dots, P_n)z$. Voter i *controls an alternative* x if and only if for all z he controls the choices (x, z) and (z, x).

Example. If the pair $\{1, 2\}$ is inseparable, then at least under the social welfare function studied by Kalai and Ritz, some voter will control every alternative except 1, 2.

Our main result will be that either the domain has size $o(m!)$ or some voter controls an arbitrarily large proportion of the alternatives.

Our proofs will involve mainly the statistics of permutations, considered as linear orders. A triple $\{x, y, z\}$ is said to be *free* in a domain D if each of the six orderings of x, y, z occurs in some permutation of D. Otherwise it is said to be *restricted.*

We sometimes use the inequality $(1-x)^y < e^{-xy}$ for $0 < x < 1$, $y > 0$ (take logarithms of both sides).

Lemma 1. *Let a be positive. For each m, choose any integer i such that $am < 2i < m/2$.*

Consider the sets of pairs $S_0 = \{(1, 2), (3, 4),\dots,(2i-1, 2i)\}$ [illegible] *probability that in a random permutation of* $\{1, 2, \ldots, m\}$ *no pair of* S_0 *occurs in order and precedes* m, *tends to zero, as* $m \to \infty$.

Proof. Let $S_1 = \{2i+1, 2i+2, \ldots, m-1\}$. Then the probability that no more than $\lfloor m^{2/3}+1 \rfloor$ elements of S_1 precede m is

$$\frac{\lfloor m^{2/3}+2 \rfloor}{|S_1+1|} \to 0.$$

To see this, note that m is equally likely to be in any position relative to the elements of S_1.

Now we consider any permutation on $S_1 \cup \{m\}$ where at least $\lfloor m^{2/3}+2 \rfloor$ elements of S_1 do precede m. We insert the numbers $(1, 2)$, $(3, 4), \ldots, (2i-1, 2i)$ into the permutation two by two. We need to do this so that both elements do not precede m and occur in order. Suppose we have inserted all pairs prior to $2j-1, 2j$. The probability that a random insertion of $2j-1$, $2j$ will put them in order, prior to m

$$\frac{1}{2}\,\frac{(u+1)(u+2)}{(v+1)(v+2)},$$

where $u \geqslant \lfloor m^{2/3}+2 \rfloor$ is the number of elements currently preceding m and $v \leqslant m-2$ is the total number of elements already arranged.

$$\frac{\binom{u+1}{2}}{\binom{v}{2}} \geqslant \frac{1}{2}\,\frac{m^{4/3}}{m^2}.$$

Thus the probability that a random insertion will be all right is at most

$$1 - \tfrac{1}{2} m^{-2/3}.$$

This is the conditional probability that the arrangement will still be all right, given that the previous arrangement was a certain acceptable arrangement. The conditional probability that the arrangement will still be all right, given the previous arrangement was all right will be a convex combination of these conditional probabilities for specific prior arrangements. If we multiply these conditional probabilities we have at most

$$\begin{aligned}(1 - \tfrac{1}{2} m^{-2/3})^{l} &\leqslant (1 - \tfrac{1}{2} m^{-2/3})^{\alpha m/2}\\ &\leqslant (e^{-\alpha m^{1/3}/4}) = o(1).\end{aligned}$$

This completes the proof.

LEMMA 2. *Let notation be as in Lemma* 1. *The probability that in a random permutation of* $\{1, 2, \ldots, m\}$ *no pair of* S_0 *occurs in order and* m *occurs between the two elements of the pair, tends to zero as* m *tends to infinity.*

Proof. The probability that either no more than $[m^{2/3}]$ elements of S_1 precede m or that at most $[m^{2/3}]$ elements of S_1 follow m is at most $2[m^{2/3}+1]/|S_1+1|$ (use the argument of Lemma 1 for each case together with $\operatorname{pr}(A \cup B) \leqslant \operatorname{pr}(A) + \operatorname{pr}(B)$ for any events A, B). This tends to zero. Thus we may suppose that at least $[m^{2/3}+1]$ elements of S_1 precede m and at least $[m^{2/3}+1]$ elements of S_1 follow m. Suppose we have an arrangement of $\{m\} \cup S_1 \cup \{1, 2, \ldots, 2j-2\}$ in which m has at least $[m^{2/3}+1]$ predecessors and at least $[m^{2/3}+1]$ successors, in S_1. For a random insertion of $2j-1$, $2j$ the probability that they will occur in order with m in between is

$$\frac{(u+1)(w+1)}{(v+1)(v+2)},$$

where u is the current number of predecessors of m, w is the current number of successors of m, and v is the current number of elements already arranged. We have $u, w \geqslant [m^{2/3}+1]$, $v \leqslant m-2$, so this quantity is at least

$$\frac{m^{4/3}}{m^2} = m^{-2/3}.$$

Thus the probability the insertion will be all right is at most $1 - m^{-2/3}$. If we multiply conditional probabilities as in Lemma 1 we have at most

$$(1 - m^{-2/3})^{\alpha m/2} \to 0.$$

This proves the lemma.

THEOREM 3. *Let* $m_l \to \infty$, *let* $0 < \beta < 1$, *and* $D(m_l)$ *be a domain of permutations in which at least* $\beta\binom{m_l}{2}$ *pairs are contained in restricted triples. Then*

$$\frac{|D(m_l)|}{m_l!} \to 0.$$

Proof. Suppose not. We may assume that

$$\lim \frac{|D(m_l)|}{m_l} = \delta > 0$$

by possibly taking a subsequence of (m_l). First suppose that the number of disjoint triples $\{x, y, z\}$ which are restricted in $D(m_l)$ is not bounded. If there are c disjoint restricted triples the maximum number of permutations in $D(m_l)$ is $(5/6)^c\, m_l!$. To see this, say that two permutations (not necessarily in D) are equivalent if they differ by a substitution which replaces each member of one of the c triples by some member of the same triple, and is the identity on the other elements. Each equivalence class has 6^c elements and at most 5^c can belong in $D(m_l)$.

If $c \to \infty$ we would have

$$\lim(5/6)^c > 0,$$

which is false.

So there is a bound c on the number of disjoint restricted triples. In each $D(m_l)$ choose a set of disjoint restricted triples of maximum size. Then every restricted triple involves one of the at most $3c$ elements in these triples. (Otherwise we would have an additional restricted triple).

The number of restricted triples is at least $\frac{1}{3}\beta\binom{m_l}{2}$ since every triple has 3 pairs and there are at least $\beta\binom{m_l}{2}$ restricted pairs. So some element w is contained in at least $(\beta/9c)\binom{m_l}{2}$ restricted triples, by the preceding paragraph.

In at least a third of these the forbidden ordering will put w in one of these positions: (i) first, (ii) middle, (iii) last.

Case 1. For an infinite subsequence of $D(m_l)$ a set T of at least $(\beta/27c)\binom{m_l}{2}$ triples involving w have a restriction such that w is last in the forbidden ordering.

We assert that we can choose a subset U of these triples having at least $\beta(m_l - 1)/108c$ members such that the intersection of any two triples in U is $\{w\}$. We will signify the forbidden ordering of such a triple as xyw. Let U be a maximal subset such that the pairs $\{x, y\}$ are disjoint. If this subset has u members then every triple $\{w, x, y\}$ contains one of the $2u$ members other than w, in T. Else U would not be maximal. Thus every triple in the set U can be obtained by adding to w one of the $2u$ elements and then adding a remaining element which can be chosen in at most m_l ways. Thus

$$2um_l \geqslant \frac{\beta}{27c}\binom{m_l - 1}{2}.$$

This implies the statement about $u = |U|$.

We assert that the number of permutations which satisfy all triple restrictions holding on $D(m_l)$ for triples in U, tends to 0. This is a consequence of Lemma 1, if we relabel the elements so that U consists of the triples $\{1, 2, m\}$, $\{3, 4, m\}$, $\{2u-1, 2u, m\}$, and take $\alpha = \beta/217c$. If $2u \geqslant m_l/2$, we take a new set U' by including only

$$\left[\frac{m_l - 1}{4}\right]$$

of the triples of U.

Case 2. For an infinite subsequence of the $D(m_l)$ a set T of at least $(\beta/27c)\binom{m_l}{2}$ triples involving w have a restriction such that w is first in the forbidden ordering.

This case is symmetrical to Case 1.

Case 3. For an infinite subsequence of the $D(m_l)$ a set T of at least $(\beta/27c)\binom{m_l}{2}$ triples involving w have a restriction such that w is in the middle in the middle in the forbidden ordering.

This case follows from Lemma 2 by reasoning similar to that in Case 1.

This completes the proof.

DEFINITION. A pair x, y of alternatives is called a *trivial pair* for a domain D if either (1) for all $P \in D$, xPy, or (2) for all $P \in D$, yPx.

THEOREM 4. *Let* $\varepsilon > 0$. *Suppose* $D(m_l)$ *be a sequence of domains admitting a binary, Pareto, nondictatorial social welfare function in which no individual controls more than* $(1-\varepsilon)\, m_l$ *alternatives. Then*

$$\frac{|D(m_l)|}{m_l!} \to 0.$$

Proof. Suppose the hypotheses of this theorem are true, but that

$$\frac{|D(m_l)|}{m_l!}$$

does not converge to zero. By Theorem 3, for any fixed $\beta > 0$, there must be at least $(1-\beta)\binom{m_l}{2}$ pairs (x, y) which are not contained in any restricted triple.

Suppose $\{x, y\}$ is a pair of alternatives not contained in a restricted triple. Then for all z distinct from x, y the triple $\{x, y, z\}$ is free and so the social welfare function restricted to this triple has a dictator, say voter i. The dictator must be constant over z, since he controls the choice (x, y) and the pair (x, y) is not trivial, being contained in a free triple. It follows that this dictator controls alternatives x, y.

Suppose (z, w) is another pair of alternatives not contained in any restricted triple. Then some voter j controls alternatives z, w. But both voters

i and j control the choice x, z. And since x, y is not contained in a restricted triple, this pair is nontrivial. So voter i and vector j are the same person. Thus some voter controls all alternatives in all pairs not contained in restricted triple. Thus if the union of the $(1-\beta)\binom{m_l}{2}$ pairs has v elements, at least v alternatives will be controlled by some voter. But

$$\binom{v}{2} \geqslant (1-\beta)\binom{m_l}{2}.$$

Thus $v^2 \geqslant (1-\beta)(m_l^2 - m_l)$, $v \geqslant (1-\beta)^{1/2} m_l(1-1/m_l)^{1/2}$ and by choice of β we can make $v > (1-\varepsilon)\, m_l$ for large m_l. This contradicts the hypothesis of the theorem.

This completes the proof.

THEOREM 5. *There exists a domain of no more than*

$$\frac{3\log(m-2)}{\log 1.2} + 1$$

permutations in which every triple is free, for any m.

Proof. Suppose we choose one-by-one, k permutations independently and at random. Suppose we let D be the resulting set of permutations. Let N be the number of free triples. Let X_T be the random variable which is 1 or 0 according as the triple T is or not free. Then $E(X_T) = \mathrm{pr}(T \text{ is free}) = 1 - \mathrm{pr}(T$ is not free in $D)$. For each of the six orderings of T, the probability that ordering is forbidden is $(5/6)^k$, that is each permutation of D can be chosen in $(5/6)m!$ ways if the forbidden triple does not occur.

Thus

$$\begin{aligned} E(N) &= \sum_T E(X_T) \\ &\geqslant \sum_T (1-6(5/6)^k) \\ &\geqslant \binom{m}{3}(1-6(5/6)^k). \end{aligned}$$

If

$$6(5/6)^k \binom{m}{3} < 1$$

then

$$E(N) > \binom{m}{3} - 1.$$

So at least one choice of the k permutations has $N = \binom{m}{3}$, i.e., all triples free. But if

$$k \geqslant \frac{3 \log m}{\log 1.2}$$

then

$$6(5/6)^k \binom{m}{3} \leqslant 6(m)^{-3} \binom{m}{3} < 1.$$

REFERENCES

1. E. KALAI AND E. MULLER, Characterization of domains admitting nondictatorial social welfare functions and nonmanipulable voting procedures, *J. Econ. Theory* **14** (1977), 457–469.
2. E. KALAI AND Z. RITZ, "An Extended Single Peak Condition in Social Choice," *Discussion Paper No. 325*, Center for Mathematical Studies in Economics and Management Sciences, Northwestern University, May, 1978.
3. K. H. KIM AND F. W. ROUSH, "Introduction to Mathematical Theories of Social Consensus," Dekker, New York, 1980.

JOURNAL OF ECONOMIC THEORY 24, 48–69 (1981)

On the Informational Size of Message Spaces for Resource Allocation Processes in Economies with Public Goods

FUMITAKA SATO

Department of Economics, Toyama University, 3190 Gofuku, Toyama 930, Japan

Received November 7, 1978; revised April 29, 1980

1. INTRODUCTION

The main purpose of this paper is to investigate the informational requirements of resource allocation processes in economies with public goods. The methods and results of Hurwicz [9], Mount and Reiter [14], Walker [17], and Osana [15] will be extended to show that the Lindahl process uses the smallest message space among a certain class of resource allocation processes that are *privacy preserving* and *non-wasteful*. In other words, the Lindahl process will be shown in a certain sense to be informationally the most efficient process for allocating public goods.

The question of informational properties of economic systems has been noted by some economists (e.g., F. A. Hayek, O. Lange, etc.) in the context of the debate over the feasibility of central planning.[1] Formal study of this question was initiated by Hurwicz [7] in 1959. Hurwicz gave an explicit formalization of the resource allocation process as an iterative procedure of information processing and communication. Inquiries as to the informational properties of resource allocation processes have been practiced on the basis of this formulation. Hurwicz [9], restricting message spaces to be Euclidean, proved that the dimensionality of the message space for the competitive process is minimal among message spaces for *privacy preserving* and *non-wasteful* resource allocation processes on a certain class of economies with private goods. Mount and Reiter [14] introduced a new concept of the informational size of topological spaces in order to prove a result similar to the one in Hurwicz [9], without the restriction to Euclidean spaces. Later, Walker [17] showed that the result of Mount and Reiter was true with a more natural definition of the size of message spaces than the one used in [14], and that this new definition is the strongest for which the result could be true. Osana [15] extended these results to a wider class of resource

---

[1] See [4 and 22], for instance.

0022-0531/81/010048-22$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

allocation processes. All of these results confirm the intuitive notion that the perfectly competitive process, which might be considered the most decentralized process, is informationally the most efficient one. In this paper, these methods and results will be extended and applied to economies with public goods to show that the Lindahl process is similarly the most efficient.

Our attention will be confined to *privacy preserving* and *non-wasteful* resource allocation processes. The privacy aspect of processes means that the information about the environment is initially dispersed through the economy. This would be the minimal requirement for the concept of informational decentralization. The imposition of *non-wastefulness* is important since the informational properties of resource allocation processes should be investigated in connection with their performance characteristics. The question addressed in this paper is, thus, as follows: In economies with public goods, what is the minimal informational size of message spaces for resource allocation processes that are *privacy preserving* and *non-wasteful*? To study this question, according to Mount and Reiter ([14]), the resource allocation process is reformulated as a composition of a message correspondence which associates equilibrium messages with environments and an outcome function which yields joint actions according to equilibrium messages. The message space is, thus, meant to be the space of equilibrium messages.

There might be a question as to the relationships between the incentive properties and the informational requirements of the resource allocation processes. This question is not considered in this paper.[2]

In the next section, a few concepts concerning the resource allocation process will be defined. In Section 3, a concept that characterizes the informational size of message spaces will be introduced, and some preliminary propositions will be proved. In Section 4, these propositions will be applied to the case of resource allocation processes for economies with public goods. Lemma 3 of this section will play the central role in finding a lower bound of the informational size of message spaces for resource allocation processes. It would be worth noting that the similar injectiveness results (or single-valuedness lemmas) are also the keys to the proofs of other results in this research area described above.

2. Resource Allocation Processes

We will consider economies with N agents. The agents are indexed by i ($i = 1, \ldots, N$). Let e_i denote the characteristic (consumption possibility, preference, endowment, production technology, etc.) of the ith agent. And let $\mathscr{E}_i$ denote the class of all possible characteristics of the ith agent. The class of all possible environments is given by the Cartesian product $\mathscr{E} = \mathscr{E} \times \cdots \times \mathscr{E}_N$.[3] Let Z denote the set of joint actions (exchanges, production, etc.).

---

[2] An example could be given which might suggest that the process by Groves and Ledyard [3] is "informationally the most efficient" over a certain class of environments.

A resource allocation process could be said to be a model describing how decisions as to joint actions are taken for a given environment. Formally, a resource allocation process is summarized by a correspondence[4]

$$f: \mathscr{E} \rightarrow Z$$

from the class of environments to the set of joint actions, called a *performance correspondence.* In general, some communications among the agents will be necessary for a consistent joint action to be taken, since the knowledge of the environment is supposed to be dispersed through the economy. Hence, the process is decomposed into two stages. In the first stage, messages are exchanged iteratively to obtain equilibrium messages. In the second stage, a joint action is taken on the basis of the equilibrium message only.[5] This can be expressed formally by saying that the performance correspondence f is a composition of a *message correspondence*

$$\mu: \mathscr{E} \rightarrow M$$

which determines the quilibrium messages for the environment $e = (e_1, \ldots, e_N) \in \mathscr{E}$, and an *outcome function*

$$h: M \rightarrow Z,$$

with $f = h \circ \mu$. The (topological) space M will be called a *message space.* In this formulation, only equilibrium messages are considered. Since any space in which the iterative communication takes place must at least include the equilibrium messages, the size of the equilibrium message space can be considered a lower bound on that of the message space of a resource allocation process.

Remark. The basic conceptual framework of the resource allocation process has been given by Hurwicz [7], where the process has been formulated as an iterative process of information processing and communication. The iterative process of communication is described by *response functions* of the agents. The response function of each agent determines messages to be sent by the agent as a function of the previously received messages and of the economic environments. After some stages of iteration are carried out, equilibrium (or stationary) messages may be obtained. Then, a joint action is taken on the basis of the equilibrium messages. Later, Mount and Reiter [14] have reformulated the notion of the resource allocation process to investigate the informational requirements of the process. In this paper, we have adopted the form presented by Mount and Reiter. The relationship of the formulation adopted here and that of Hurwicz [7] could be found in [14].

---

[3] If environments are non-decomposable, the class $\mathscr{E}$ of environments may not be given as a Cartesian product of each agent's class of characteristics. In this paper, only the case where $\mathscr{E}$ is given by a product will be considered, although this restriction is unnecessary.

[4] Let S and T be nonempty sets. A correspondence ϕ from S to T ($\phi: S \rightarrow T$) is a rule that associates with each element $x \in S$ a (possibly empty) subset $\phi(x)$ of T.

[5] Outcome functions (defined below) are assumed "nonparametric" in the sense of Hurwicz [8].

Hurwicz [8] has presented examples based on the Peano curve to show that there exist encoding procedures by which a simple message conveys unlimited amount of information. As Hurwicz notes, such procedures use, in a certain sense, "tricks" not inherent to the resource allocation process. This suggests that in order to study the "genuine" (i.e., without "tricks" or inherent to the resource allocation process) informational requirements of the allocation process we must impose a regularity condition upon message correspondences. We will impose a (topological) regularity condition upon message correspondences, which is defined as follows.

DEFINITION 1. Let S and T be topological spaces. The correspondence $\phi: S \to T$ is said to be *locally threaded* if and only if for every $x \in S$ there exist a neighborhood U of x and a continuous mapping $s: U \to T$ such that $s(u) \in \phi(u)$ for all $u \in U$.

Imposing this regularity condition, we will define the resource allocation process as follows.

DEFINITION 2. Let $\bar{\mathscr{E}}\ (\subset \mathscr{E})$ and M be topological spaces, let $\mu: \bar{\mathscr{E}} \to M$ be a correspondence, and let $h: M \to Z$ be a mapping. Then the triplet $\Pi = (M, \mu, h)$ is said to be a *resource allocation process* on $\bar{\mathscr{E}}$ if and only if the correspondence μ is locally threaded on $\bar{\mathscr{E}}$.

Remark. Hurwicz [9] adopted a regularity condition such that the inverse of the message correspondence had a Lipschitzian selection (see [9, p. 416]). The regularity condition imposed in this paper is the one adopted by Mount and Reiter [14]. Each of these regularity conditions excludes the possibility that the inverse of the message correspondence is the Peano function, though they are mutually independent.

Generally, the knowledge about the environment is supposed to be dispersed through the economy, and hence the initial concentration of information to some agents will be excluded. This is considered the minimal requirement for the concept of informational decentralization. Following Mount and Reiter, this requirement can be defined as follows.

DEFINITION 3. Let $\mathscr{E} = \mathscr{E}_1 \times \cdots \times \mathscr{E}_N$ be a product space. The correspondence $\mu: \mathscr{E} \to M$ is said to be a *coordinate correspondence* on $\mathscr{E}$ if and only if there exists a correspondence $\mu^i: \mathscr{E}_i \to M_i$ for each i ($i = 1,\ldots, N$) such that for all $e = (e_1,\ldots, e_N) \in \mathscr{E}$,

$$\mu(e) = \bigcap_{i=1}^{N} \mu^i(e_i),$$

where $\bigcap_{i=1}^{N} M_i \supset M$.

DEFINITION 4. A resource allocation process $\Pi = (M, \mu, h)$ is said to be *privacy preserving* on $\mathscr{E}$ if and only if the message correspondence μ is a coordinate correspondence on $\mathscr{E}$.

When a process is privacy preserving, each agent's messages are dependent on the environments only through the characteristics of the agent and the agent does not need to know the characteristics of the other agents.

Informational properties of a process should be examined in connection with its other properties. In this paper, informational properties will be investigated for a class of resource allocation processes which yield Pareto optimal joint actions. Let $\mathfrak{P}(e) \subset Z$ be a set of Pareto optimal joint actions for $e \in \mathscr{E}$. Then, for any subclass $\mathscr{E}$ of $\mathscr{E}$, an *optimality correspondence* $\mathfrak{P}: \mathscr{E} \to Z$ will be defined as a correspondence which associates (possibly empty) Pareto optimal actions $\mathfrak{P}(e)$ with each $e \in \mathscr{E}$.

DEFINITION 5. A resource allocation process $\Pi = (M, \mu, h)$ is said to be *non-wasteful on $\mathscr{E}$ with respect to $\mathfrak{P}$* if and only if $h \circ \mu(e) \subset \mathfrak{P}(e)$ for every e of $\mathscr{E}$, where $h \circ \mu(e) = \bigcup_{m \in \mu(e)} h(m)$.

If $\mathfrak{P}(e)$ is the set of *all* Pareto optimal joint actions for each e of $\mathscr{E}$, and if the process Π is non-wasteful on $\mathscr{E}$ with respect to $\mathfrak{P}$, then Π is said simply to be non-wasteful on $\mathscr{E}$. In other words, a process is non-wasteful if its outcomes, if any, are Pareto optimal for each environment e of a given class of environments.

Remark. Some concrete examples will serve to illustrate these concepts. It is well known that the perfectly competitive process is non-wasteful over a certain class of "classical" environments (i.e., environments free from indivisibilities, non-convexities, externalities, or public goods).[6] Hurwicz's *greed process* (see [7]) is non-wasteful over a wide class of decomposable environments. The Lindahl process is non-wasteful over a certain class of environments with public goods.[7] It is easy to verify that they are privacy preserving resource allocation processes on certain (topological) spaces of environments.

---

[6] See Koopmans [11] for the non-wastefulness of the (perfectly) competitive process.

[7] See Milleron [13].

In the following sections, the minimal informational size of message spaces will be investigated for resource allocation processes which are privacy preserving and non-wasteful over a (topological) space of environments with public goods.

3. INFORMATIONAL SIZE OF MESSAGE SPACES

The notion of informational size could be considered a concept which characterizes the relative sizes of topological spaces that are used to convey information in resource allocation processes. It would be natural to consider that a space, say S, has more information than the other space T whenever S is topologically "larger" than T. This suggests the following definition.

DEFINITION 6. Let S and T be topological spaces. Then, $S \geqslant_F T$ if and only if T can be embedded homeomorphically in S, i.e., there is a subspace S' of S which is homeomorphic to T. (The space S is said to have as much information as the space T by the Fréchet ordering.)[8]

Remark. A concept of informational size was introduced by Mount and Reiter [14] in proving the "informational efficiency" of the competitive process. Walker [17] examined some simpler and more natural concepts of the size of topological spaces and showed that the concept adopted in this paper was more appropriate for the Mount–Reiter analysis than the one used in [14]. As Walker points out, if the space T is a finite-dimensional Euclidean space, then the Fréchet ordering $\geqslant_F$ is equivalent to a variant (which he denotes $\geqslant_S$) of the Mount–Reiter ordering. Several of the results to follow are therefore true for $\geqslant_S$ as well as for $\geqslant_F$.

Two lemmata on the informational size will be proved. We need the following definition.

DEFINITION 7. Let $\phi: T \to S$ be a correspondence. Then, ϕ is said to be *injective* if and only if $\phi(t) \cap \phi(t') \neq \varnothing$ implies $t = t'$ for any $t, t' \in T$.

The next lemma is an extension of Proposition 3 of Calsamiglia [2].

LEMMA 1. *Let T be a space which is homeomorphic to a finite-dimensional Euclidean space and S be a Hausdorff space. If there is a locally threaded injective correspondence $\phi: T \to S$, then $S \geqslant_F T$.*

---

[8] This preordering $\geqslant_F$ is the concept introduced by Fréchet to define the "dimension type."

Proof. Without loss of generality, T will be supposed to be a finite-dimensional Euclidean space, since the preordering $\geqslant_F$ is a topological invariant. Choose a point t in T. Since the correspondence ϕ is locally threaded, there exist a neighborhood V of t and a continuous mapping $s\colon V \to S$ such that $s(v) \in \phi(v)$ for every $v \in V$. This mapping s is an injection from V into S, since ϕ is an injective correspondence.

Since T is Euclidean, there exists a compact set $\bar{V} \subset V$ with nonempty interior. The restriction $s|_{\bar{V}}\colon \bar{V} \to s(\bar{V})$ of s is a continuous bijection from a compact set $\bar{V}$ onto a Hausdorff space $s(\bar{V})$. Hence it follows that the restriction $s|_{\bar{V}}$ is a homeomorphism (see Kelley [10, Theorem 5.8, p. 141]). Choose an open ball $V^\circ \subset \bar{V}$ in T. Then, V° and $s(V^\circ)$ are homeomorphic by a homeomorphism $s|_{V^\circ}\colon V^\circ \to s(V^\circ)$. This, together with the fact that the Euclidean space T is homeomorphic to its open ball V°, implies that T is homeomorphic to $s(V^\circ) \subset S$. Hence, it follows that $S \geqslant_F T$. Q.E.D.

In Lemma 1, the condition that S be Hausdorff cannot be dispensed with. The following example, which is intimately related to Example 1 of Walker [17], will serve to illustrate this. Let T be the open interval $(0, 1)$ of the real line R with the usual topology, and S be the interval $(0, 1)$ with the topology consisting of the complements of finite subsets and S itself (i.e., the proper closed subsets of S are precisely the finite subsets). The space S is not Hausdorff (though it is a T_1-space). Obviously, there is a locally threaded injective correspondence from T to S (e.g., the identity map on $(0, 1)$). However, there is no subspace of S that is homeomorphic to T. Hence, $S \geqslant_F T$ fails to hold. When we remove the condition that S be Hausdorff, the following lemma holds, where the inverse of the locally threaded correspondence is supposed to be continuous.[9]

LEMMA 2. *Let S and T be topological spaces, and T be homeomorphic to a finite-dimensional Euclidean space. If there is a locally threaded injective correspondence $\phi\colon T \to S$ such that the inverse $\phi^{-1}\colon S \to T$ is a continuous mapping on $\phi(T)$, then it holds that $S \geqslant_F T$.*

Proof. As in the proof of Lemma 1, T can be supposed to be a finite-dimensional Euclidean space. Choose a point t in T. Since ϕ is locally threaded, there exist a neighborhood V of t and a continuous mapping $s\colon V \to S$ such that $s(v) \in \phi(v)$ for all $v \in V$. Since V is a neighborhood of t in a Euclidean space T, there is an open ball $V^\circ \subset V$ in T with center t. Let $U^\circ = s(V^\circ)$, then the restriction $\phi^{-1}|_{U^\circ}$ of the inverse ϕ^{-1} to U° is a continuous mapping by the supposition in the lemma, and obviously it is the

---

[9] The (lower) inverse of the correspondence $\phi\colon T \to S$ is a correspondence $\phi^{-1}\colon S \to T$ defined by $\phi^{-1}(x) = \{t \in T \mid x \in \phi(t)\}$ for each $x \in S$. Note that ϕ^{-1} is a mapping on $\phi(T)$ if and only if ϕ is injective on T.

the inverse of the restriction $s|_{V^0}$ of the local thread s to V^0. Hence, V^0 is homeomorphic to U^0. Since the Euclidean space T is homeomorphic to its open ball V^0, T is homeomorphic to $U^0 \subset S$. This means that $S \geqslant_F T$.

Q.E.D.

By using the above lemmata, we can obtain two propositions on the informational size of message spaces for resource allocation processes. Proposition 1 (Proposition 2) is an immediate consequence of Lemma 1 (Lemma 2 respectively).

PROPOSITION 1. *Let* $\mathscr{E}$ *be a topological space of environments and* $\Pi = (M, \mu, h)$ *be a resource allocation process on* $\mathscr{E}$. *Suppose that there is a subspace* E *of* $\mathscr{E}$ *such that* (1) *the subspace* E *is homeomorphic to a* K*-dimensional Euclidean space* R^K *for some finite integer* K, *and* (2) *the message correspondence* μ *is injective on* E. *Then, if* M *is a Hausdorff space, it holds that* $M \geqslant_F R^K$.

Proof. Since the message correspondence μ is locally threaded on $\mathscr{E}$, its restriction $\mu|_E \colon E \to M$ is locally threaded, and is injective by supposition. Hence, by Lemma 1, it follows that $M \geqslant_F E$. This implies that $M \geqslant_F R^K$.

Q.E.D.

PROPOSITION 2. *Let* $\mathscr{E}$ *be a topological space of environments and* $\Pi = (M, \mu, h)$ *be a resource allocation process on* $\mathscr{E}$. *Suppose that there is a subspace* E *of* $\mathscr{E}$ *such that* (1) *the subspace* E *is homeomorphic to a* K*-dimensional Euclidean space* R^K *for some finite integer* K, *and* (2) *the restriction* $\mu|_F$ *of the message correspondence* μ *to* E *is injective and its inverse* $\mu|_E^{-1}$ *is a continuous mapping on* $\mu(E)$. *Then it holds that* $M \geqslant_F R^K$.

The simple proof of Proposition 2 is omitted since it is evident from Lemma 2 above.

These propositions will be applied to a class of resource allocation processes for environments with public goods in the next section.

Remark. Proposition 2 parallels the method used in Osana [15] in that it utilizes the continuity property of the inverse of the message correspondence. Proposition 1, in contrast with that, does not suppose the continuity of the inverse of the message correspondence, assuming the message space to be Hausdorff.[10]

---

[10] Osana [15] considered a case where the inverse of the message correspondence was lower hemi-continuous. However, the injectiveness of the correspondence implies its inverse is a mapping. Hence, the lower hemi-continuity of the inverse means its continuity.

4. ENVIRONMENTS WITH PUBLIC GOODS

4.1. *Economies with Public Goods*

In this section, the environments will be specified as the ones containing public goods. We will assume that there are l private goods (indexed $j = 1, \dots, l$) and q public goods (indexed $k = 1, \dots, q$), and the commodity space will be assumed to be the $l + q$-dimensional Euclidean space R^{l+q} or equivalently the s-dimensional Euclidean space R^s, where $s = l + q$.

There are n consumers (indexed $i = 1, \dots, n$) and a single producer (denoted by the index N), and hence there are $n + 1$ $(= N)$ economic agents. Each agent is characterized by his *endowments* of the private goods, his *consumption possibility set*, and by his *preference relation* on his consumption possibility set, or by his *production possibility set*. For the ith consumer ($i = 1, \dots, n$), let $X_i \subset R^s$ be his consumption possibility set, $\gtrsim_i$ be a reflexive binary relation on X_i which represents his preferences, and $w_i \in R^l$ be his endowments of the private goods. For the producer ($i = N$), let $Y \subset R^s$ be his production possibility set, and $w_N \in R^l$ be his endowments of the private goods.[11] Initial endowments of the public goods are assumed to be zero.

The *characteristic* of the ith consumer ($i = 1, \dots, n$) is defined by

$$e_i = (X_i, \gtrsim_i, w_i),$$

and that of the producer is defined by a pair

$$e_N = (Y, w_N).$$

The class of all possible characteristics of the ith agent is denoted by $\mathscr{E}_i$ ($i = 1, \dots, n, N$). The *environment* is defined as

$$\begin{aligned} e &= ((e_i)_{i=1}^n, e_N) \\ &= (e_1, \dots, e_n, e_N), \end{aligned}$$

and the class of all possible environments $\mathscr{E}$ is supposed to be the Cartesian product

$$\mathscr{E} = \prod_{i=1}^{n} \mathscr{E}_i \times \mathscr{E}_N.$$

---

[11] Usually, a producer is supposed to have no endowments. If a production possibility set Y_w is defined as $Y_w = Y + \{(w_N, 0)\}$, then the producer with the production possibility set Y_w and without endowments can be identified with a producer whose production possibility set is Y and endowments are w_N. Hence, the formulation in the text is virtually identical with the usual formalization of the producer. The formulation in the text was taken for convenience' sake.

Remark. The graph $\{(\eta, v) \in X \times X \mid \eta \succsim v\}$ of the relation $\succsim$ is called the *preference set* associated with the consumption set X and the preference relation $\succsim$. We can define the preferences $(X_i, \succsim_i)$ of the ith consumer by a preference set. Then, for instance, we may define the class of characteristics of the ith consumer as follows: Let $\mathscr{P}_i$ be the class of all possible preference sets (for example all closed preference sets in $R^s \times R^s$) of the ith consumer ($i = 1,\ldots, n$), and $W_i \subset R^l$ be the set of all possible endowments of the private goods of the ith consumer. Then, the class $\mathscr{E}_i$ is defined as $\mathscr{E}_i = \mathscr{P}_i \times W_i$. We can endow the class $\mathscr{E}_i$ with a suitable topology. For instance, if $\mathscr{P}_i$ is the class of all the closed preference sets in $R^s \times R^s$, we can endow it with the topology of closed convergence (see [5] or [6]). And we can give W_i the usual relative topology of R^s. Then, we can endow $\mathscr{E}_i$ with the product topology. Similarly, the class $\mathscr{E}_N$ can be endowed with a suitable topology.

Let $z_i = (z_{1i},\ldots, z_{li})$ denote the *net exchange* of the private goods by the ith consumer. If the sign of z_{ji} is positive, the ith consumer is a net demander of the jth private good. Let $\zeta = (\zeta_1,\ldots, \zeta_l)$ denote the (negative) *net exchange* of the private goods by the producer. Contrary to the case of the consumers, if ζ_j is positive, the producer is a net supplier of the jth private good. Let $x_i = (x_{1i},\ldots, x_{qi})$ denote a *consumption vector* of the public goods by the ith consumer, and let $\xi = (\xi_1,\ldots, \xi_q)$ denote a *production vector* of the public goods. The vector $a_i = (z_i, x_i) \in R^s$ will be called an *action* of the ith consumer ($i = 1,\ldots, n$), and the vector $a_N = (\zeta, \xi) \in R^s$ will be called an *action* of the producer. Then,

$$\begin{aligned} a &= ((a_i)_{i=1}^n, a_N) \\ &= (((z_i, x_i))_{i=1}^n, (\zeta, \xi)) \end{aligned}$$

will be called a *joint action*. A joint action a is said to be *consistent* if and only if

$$\sum_{i=1}^{n} z_i = \zeta, \tag{1}$$

and

$$x_i = \xi, \qquad i = 1,\ldots, n, \tag{2}$$

i.e., the total net demand for the private goods by consumers is equal to the net supply by the producer, and the consumption of the public goods by each consumer is equal to the production by the producer. Let Z denote the set of all consistent actions. A joint action $a = (((z_i, x_i))_{i=1}^n, (\zeta, \xi))$ is said to be *feasible* for an environment $e = (((X_i, \succsim_i, w_i))_{i=1}^n, (Y, w_N))$ if and only if

$$a \in Z, \tag{3}$$

$$(\zeta, \xi) \in Y + \{(w_N, 0)\}, \tag{4}$$

and

$$(z_i + w_i, x_i) \in X_i, \qquad i = 1, \dots, n, \tag{5}$$

i.e., it is consistent and possible for all agents.

A feasible joint action $a = (((z_i, x_i))_{i=1}^n, (\zeta, \xi))$ is said to be *Pareto optimal* for an environment e if and only if there does not exist a feasible joint action $a' = (((z'_i, x'_i))_{i=1}^n, (\zeta', \xi'))$ such that (i)

$$(z'_i + w_i, x'_i) \succsim_i (z_i + w_i, x_i) \tag{6}$$

for all consumers ($i = 1, \dots, n$), and (ii)

$$\text{not } (z_\iota + w_\iota, x_\iota) \succsim_\iota (z'_\iota + w_\iota, x'_\iota) \tag{7}$$

for some consumer ι.

4.2. *Cobb–Douglas=Linear* (CDL) *Economies*

The propositions in the previous section will be applied to the environments with public goods. For this purpose, a subclass E of environments will be defined, which will be characterized by Cobb–Douglas utility functions and linear production functions.

For given parameters $B^i = (b_2^i, \dots, b_l^i; \beta_1^i, \dots, \beta_q^i)$ in the interior of the positive orthant of R^{s-1} (i.e., $B^i \in R_{++}^{s-1}$), we can define a Cobb–Douglas utility function $U^{B^i}: R_+^s \to R$ on the nonnegative orthant of the s-dimensional Euclidean space by

$$U^{B^i}(z_i, x_i) = z_{1i} \prod_{j=2}^{l} (z_{ji})^{b_j^i} \cdot \prod_{k=1}^{q} (x_{ki})^{\beta_k^i}$$

for $(z_i, x_i) \in R_+^s$. Define a preference set $\rho(B^i)$ associated with the Cobb–Douglas utility function U^{B^i} by

$$\rho(B^i) = \{(\eta, \upsilon) \mid \eta \in R_+^s,\ \upsilon \in R_+^s, \text{ and } U^{B^i}(\eta) \geqslant U^{B^i}(\upsilon)\},$$

and let P denote the class of all the preference sets associated with Cobb–Douglas utility functions U^{B^i} ($B^i \in R_{++}^{s-1}$), i.e.,

$$P = \{\rho(B^i) \mid B^i \in R_{++}^{s-1}\}.$$

Now, define the subclass E_i of characteristics of the ith consumer by

$$E_i = P \times \{0\},$$

where 0 denotes the zero vector of R^l ($i = 1, \dots, n$). Then, E_i is the class of characteristics of the ith consumer with preference relations that can be represented by Cobb–Douglas utility functions and without resource endowments of the private goods.

For given parameters $C = (c_2, \ldots, c_l; \gamma_1, \ldots, \gamma_q)$ in the interior of the positive orthant of R^{s-1}, define a production possibility set

$$Y^C = \left\{ (y, \xi) \in R^s \mid y_1 + \sum_{j=2}^{l} c_j y_j + \sum_{k=1}^{q} \gamma_k \xi_k \leqslant 0 \right\},$$

and define the subclass E_N of the characteristics of the producer by

$$E_N = \{ (Y^C, w_N) \mid C \in R^{s-1}_{++}, w_N = (\omega, 1, \ldots, 1) \in R^l, \text{and } \omega \in R_{++} \}.$$

Then, E_N is the class of characteristics of the producer which can be represented by linear production functions and positive endowments of the private goods. Note that the endowments of the private goods other than that of the first private good are fixed to unity.

Now, let $E = \prod_{i=1}^{n} E_i \times E_N$. This class E might be called *Cobb–Douglas=Linear* (CDL) *economies.*[12] A topology is introduced to the class E as follows. For each consumer ($i = 1, \ldots, n$), define a metric d^i on E_i by $d^i(\rho(B^i), \rho(B^{i\prime})) = \delta(B^i, B^{i\prime})$, where δ is the usual Euclidean distance of R^{s-1}, and define a metric d^N on E_N by $d^N((Y^C, w_N), (Y^{C'}, w'_N)) = \bar{\delta}((C, \omega), (C', \omega'))$, where $\bar{\delta}$ is the usual Euclidean distance of R^s. Let E have the product topology, which will be denoted by $\mathscr{T}$. Then, it will be clear that the topological space $(E, \mathscr{T})$ is homeomorphic to the $N(s-1)+1$-dimensional Euclidean space $R^{N(s-1)+1}$.

Remark. It is worth noting that, for instance, if a class $\mathscr{E}\ (\supset E)$ is endowed with the product topology of the topologies of closed convergence for agents' characteristics (see Hildenbrand [5] or [6]), then its relative topology for E coincides with the topology $\mathscr{T}$ given to E in the above. See [16].

Given a class $\mathscr{E}$ of environments which includes E, we will define an optimality correspondence $\mathfrak{P}: \mathscr{E} \to Z$ such that the restriction $\mathfrak{P}|_E$ associates with $e \in E$ the set $\mathfrak{P}(e)$ of all the Pareto optimal joint actions which assign strictly positive consumption to every consumer.

The next lemma is central in finding the lower bound of informational requirements of resource allocation processes.

LEMMA 3. *Let $\Pi = (M, \mu, h)$ be a resource allocation process that is privacy preserving and non-wasteful with respect to $\mathfrak{P}$ on E. Then, the message correspondence μ is injective on E.*

---

[12] This designation was suggested by a referee.

Proof. In the first place, the following notation will be defined. If $e = (e_1, \ldots, e_N)$ and $e' = (e'_1, \ldots, e'_N)$ are elements of the class E, we will denote by $e' \otimes_i e$ the element $(e_1, \ldots, e_{i-1}, e'_i, e_{i+1}, \ldots, e_N)$ for $i = 1, \ldots, N$.

Now, suppose that there is a message $m \in \mu(e) \cap \mu(e')$ for $e, e' \in E$. It will be proved that $e = e'$. Let $a = (((z_i, x_i))_{i=1}^n, (\zeta, \xi))$ be a joint action corresponding to the message m, i.e., $a = h(m)$. Since μ is a coordinate correspondence,

$$\mu(e) \cap \mu(e') = \mu(e' \otimes_i e) \cap \mu(e \otimes_i e')^{13} \tag{8}$$

for all i ($i = 1, \ldots, N$), and hence, in particular,

$$m \in \mu(e) \cap \mu(e' \otimes_i e) \tag{9}$$

for $i = 1, \ldots, N$. Since the process Π is non-wasteful with respect to $\mathfrak{P}$, the supposition that $a = h(m)$ and (9) imply that

$$a \in \mathfrak{P}(e) \cap \mathfrak{P}(e' \otimes_i e) \tag{10}$$

for $i = 1, \ldots, N$. By the usual Lagrangian method of constrained maximization, $a \in \mathfrak{P}(e)$, in turn, implies

$$b_j^i \cdot \frac{z_{1i}}{z_{ji}} = c_j, \qquad j = 2, \ldots, l; \quad i = 1, \ldots, n, \tag{11}$$

$$\sum_{i=1}^{n} \beta_k^i \cdot \frac{z_{1i}}{x_{ki}} = \gamma_k, \qquad k = 1, \ldots, q, \tag{12}$$

and

$$(\zeta_1 - \omega) + \sum_{j=2}^{l} c_j(\zeta_j - 1) + \sum_{k=1}^{q} \gamma_k \xi_k = 0.^{14} \tag{13}$$

Similarly, $a \in \mathfrak{P}(e' \otimes_i e)$ implies

$$b_j^{i\prime} \cdot \frac{z_{1i}}{z_{ji}} = c_j, \qquad j = 2, \ldots, l, \tag{14}$$

and

$$\sum_{\iota = 1, \iota \neq i}^{n} b_k^{\iota} \cdot \frac{z_{1\iota}}{x_{k\iota}} + \beta_k^{i\prime} \cdot \frac{z_{1i}}{x_{ki}} = \gamma_k, \qquad k = 1, \ldots, q, \tag{15}$$

---

[13] See Mount and Reiter [14, p. 171, Lemma 5].

[14] Note that $a \in \mathfrak{P}(e)$ implies the consumption (z_i, x_i) is positive for each consumer. See the definition of $\mathfrak{P}$ (Remark, Section 4.2).

for the ith consumer ($i = 1,\dots, n$). Then, (11), (12), (14) and (15) imply

$$B^i = B^{i\prime}, \qquad \text{for } i = 1,\dots, n, \tag{16}$$

since $z_{1i} \neq 0$. As for the producer, $a \in \mathfrak{P}(e' \otimes_N e)$ implies

$$b_j^i \cdot \frac{z_{1i}}{z_{ji}} = c_j', \qquad j = 2,\dots, l; \quad i = 1,\dots, n, \tag{17}$$

$$\sum_{i=1}^{n} \beta_k^i \cdot \frac{z_{1i}}{x_{ki}} = \gamma_k', \qquad k = 1,\dots, q, \tag{18}$$

and

$$(\zeta_1 - \omega') + \sum_{j=2}^{l} c_j'(\zeta_j - 1) + \sum_{k=1}^{q} \gamma_k' \xi_k = 0. \tag{19}$$

The equalities (11), (12), (17) and (18) together imply

$$C = C', \tag{20}$$

and (13), (19) and (20) imply

$$\omega = \omega'. \tag{21}$$

Equalities (16), (20) and (21) mean that

$$e = e'. \tag{22}$$

Consequently, the correspondence μ is injective on E. Q.E.D.

Remark. Results similar to Lemma 3 are central in proofs of informational requirements of resource allocation processes. The crucial role played by a result similar to Lemma 3 was noted by Hurwicz [9], where Hurwicz called it the *single-valuedness lemma.* The lemma states that the inverse of the message correspondence on a certain subclass of environments, say E^0, is a (single-valued) mapping. Then, Hurwicz's proof is as follows: If the (topological) space E^0 is homeomorphic to a K-dimensional Euclidean space R^K, and if the inverse of the message correspondence is a Lipschitzian function from the (Euclidean) message space M^0 to R^K, then the dimensionality of M^0 is not smaller than K. Mount and Reiter [14] and others ([17], [15]) have utilized some types of the single-valuedness lemma though their proofs are different from that of Hurwicz.

4.3. *Minimality of the Lindahl Message Space*

In this section, first, we will find a lower bound of the informational size of message spaces for resource allocation processes that are privacy preserving and non-wasteful with respect to $\mathfrak{P}$ over a (topological) space of environments which includes E as a subspace. And next, we will ascertain that the lower bound is exactly the size of the message space of the Lindahl process.

The following theorems establish a lower bound of informational size of message spaces.

THEOREM 1. *Let $\mathscr{E}$ be a topological space of environments which includes $(E, \mathscr{T})$ as a subspace. And let $\Pi = (M, \mu, h)$ be a resource allocation process on $\mathscr{E}$ that is privacy preserving and non-wasteful with respect to $\mathfrak{P}$ on E. Then, if M is a Hausdorff space, it holds that $M \geqslant_F R^K$, where $K = N(s-1)+1$.*

Proof. As was noted above, E is homeomorphic to R^K. By Lemma 3, the restriction $\mu|_E$ of the message correspondence μ to E is injective, since Π is privacy preserving and non-wasteful with respect to $\mathfrak{P}$ on E. From Proposition 1, it follows that $M \geqslant_F R^K$. Q.E.D.

COROLLARY. *Under the assumptions of Theorem* 1, *if M is a separable metric space, then* $\dim M \geqslant K$.

The following theorem is an immediate consequence of Proposition 2 and Lemma 3. The simple proof is omitted.

THEOREM 2. *Let $\mathscr{E}$ be a topological space of environments which includes $(E, \mathscr{T})$ as a subspace. And let $\Pi = (M, \mu, h)$ be a resource allocation process on $\mathscr{E}$ that is privavy preserving and non-wasteful with respect to $\mathfrak{P}$ on E. Then, if the inverse of the restriction $\mu|_E$ of the message correspondence to E is upper hemi-continuous or lower hemi-continuous on $\mu(E)$, it holds that $M \geqslant_I R^K$, where $K = N(s-1)+1$.*

The above theorems establish a lower bound of informational size of message spaces for resource allocation processes which are privacy preserving and non-wasteful (with respect to $\mathfrak{P}$) over a (topological) space of environments with public goods. It should be shown that the lower bound is actually the minimal size. It will be clear that the space E itself may be a message space whose informational size is minimal among message spaces for privacy preserving and non-wasteful resource allocation processes over E (with respect to $\mathfrak{P}$). For instance, let Id_E be the identity map on E, and $r: E \to Z$ be an arbitrary mapping with $r(e) \in \mathfrak{P}(e)$ for all $e \in E$, then (E, Id_E, r) is a non-wasteful resource allocation process with respect to $\mathfrak{P}$ defined on E. It is also a privacy preserving process. Consequently, E is a message space whose informational size is minimal among message spaces for privacy preserving and non-wasteful resource allocation process on E (with respect to $\mathfrak{P}$). This reasoning, however, cannot be applied to wider classes of environments than E. If we are to establish the minimal size of message spaces for a wider class of environments, we should follow another reasoning. For this purpose, the *Lindahl process* will be formulated, and it will be shown that the Lindahl process is a privacy preserving and non-wasteful resource allocation process (with respect to $\mathfrak{P}$) over a space of environments which includes E as a subspace, and that it uses a message space whose informational size is minimal among message spaces for resource allocation processes which are privacy preserving and non-wasteful with respect to $\mathfrak{P}$ on E.

Messages of the Lindahl process consist of (1) the prices of the private goods other than the first private good ($p = (p_2, \ldots, p_l)$), (2) proposed exchanges of the private goods other than the first private good ($z_{*i} := (z_{2i}, \ldots, z_{li})$) for each consumer ($i = 1, \ldots, n$), (3) personalized prices of the public goods ($\pi_i = (\pi_{1i}, \ldots, \pi_{qi})$) for each consumer ($i = 1, \ldots, n$), (4) proposed production of the public goods ($\xi = (\xi_1, \ldots, \xi_q)$), and (5) the information about profit of the producer (σ). Here, the first private good is chosen as *numeraire* and hence its price is identically equal to unity. Information as to the net demand for the first private good is unnecessary, since each agent's budget constraint is always binding here.

Now, we will define a *Lindahl process* for a class of environments $\mathscr{E} = \prod_{i=1}^{n} \mathscr{E}_i \times \mathscr{E}_N$ with fixed profit shares $(\theta_i)_{i=1}^{n}$ with $\theta_i > 0$ for $i = 1, \ldots, n$ and $\sum_{i=1}^{n} \theta_i = 1$. In the rest, if η and υ are vectors, $\eta \cdot \upsilon$ denotes the inner product of η and υ, and ζ_* denotes the subvector $(\zeta_2, \ldots, \zeta_l)$ of $\zeta \in R^l$.

Each consumer makes his decision as follows: He chooses net exchanges $\bar{z}_i$ of the private goods and proposed production vector $\bar{\xi}$ that maximize his preferences subject to the budget constraint $z_{1i} + \bar{p} \cdot z_{*i} + \bar{\pi}_i \cdot \xi = \bar{\theta}_i \bar{\sigma}$ under given prices $(\bar{p}, (\bar{\pi}_i)_{i=1}^{n})$ and the profit $\bar{\sigma}$, where z_{*i} denotes $(z_{2i}, \ldots, z_{li})$. Note that the budget constraint is defined by the equality here. Define the *demand correspondence* $\Psi_i \colon R^{l-1} \times R^{nq} \times R \times \mathscr{E}_i \to R^s$ by

$$\begin{aligned}\Psi_i(\bar{p}, (\bar{\pi}_i)_{i=1}^{n}, \bar{\sigma}; e_i) = \{(\bar{z}_i, \bar{\xi}) \mid (\bar{z}_i + w_i, \bar{\xi}) \succsim_i (z_i + w_i, \xi) \\ \text{for all } (z_i + w_i, \xi) \in X_i \\ \text{such that } z_{1i} + \bar{p} \cdot z_{*i} + \bar{\pi}_i \cdot \xi = \theta_i \bar{\sigma}\},\end{aligned}$$

for $i = 1, \ldots, n$. Now define a correspondence $\nu_\theta^i \colon \mathscr{E}_i \to R^K$ ($K = N(s-1) + 1$) for each consumer by

$$\begin{aligned}\nu_\theta^i(e_i) = \{(\bar{p}, (\bar{\pi}_i)_{i=1}^{n}, (\bar{z}_{*i})_{i=1}^{n}, \bar{\xi}, \bar{\sigma}) \mid \bar{p} \in R^{l-1}, \bar{\pi}_i \in R^q \text{ and} \\ \bar{z}_{*i} \in R^{l-1} \text{ for } i = 1, \ldots, n, \bar{\sigma} \in R, \text{ and} \\ (\bar{z}_{*i}, \bar{\xi}) \in \mathrm{pr}_1\, \Psi_i(\bar{p}, (\bar{\pi}_i)_{i=1}^{n}, \bar{\sigma}; e_i)\}\end{aligned}$$

for each $e_i \in \mathscr{E}_i$, where pr_1 is the projection on R^{s-1} in R^s.

The producer maximizes the profit $\sigma = \zeta_1 + \bar{p} \cdot \zeta_* + (\sum_{\iota=1}^{n} \bar{\pi}_\iota) \cdot \xi$ under given prices $(\bar{p}, (\bar{\pi}_\iota)_{\iota=1}^{n})$ subject to the constraint that actions (ζ, ξ) should belong to the production possibility set $Y + \{(w_N, 0)\}$. Define the *supply correspondence* $\Phi: R^{l-1} \times R^{nq} \times \mathscr{E}_N \to R^s$ by

$$\begin{aligned} \Phi(\bar{p}, (\bar{\pi}_\iota)_{\iota=1}^{n}; e_N) = \{(\bar{\zeta}, \bar{\xi}) \in Y + \{(w_N, 0)\} \mid (\bar{\zeta}, \bar{\xi}) &\text{ maximizes} \\ &\text{the profit } \zeta_1 + \bar{p} \cdot \zeta_* + (\textstyle\sum_{\iota=1}^{n} \bar{\pi}_\iota) \cdot \xi \\ &\text{subject to } (\zeta, \xi) \in Y + \{(w_N, 0)\}\}. \end{aligned}$$

And define a correspondence $\nu^N: \mathscr{E}_N \to R^K$ by

$$\begin{aligned} \nu^N(e_N) = \Bigg\{ (\bar{p}, (\bar{\pi}_\iota)_{\iota=1}^{n}, (\bar{z}_{*\iota})_{\iota=1}^{n}, \bar{\xi}, \bar{\sigma}) \mid \bar{p} \in R^{l-1}, &\ \bar{\pi}_\iota \in R^q \text{ and} \\ \bar{z}_{*\iota} \in R^{l-1} &\text{ for } \iota = 1,\dots, n, \\ \bar{\sigma} = \bar{\zeta}_1 + \bar{p} \cdot &\left(\sum_{\iota=1}^{n} \bar{z}_{*\iota}\right) + \left(\sum_{\iota=1}^{n} \bar{\pi}_\iota\right) \cdot \bar{\xi} \text{ with} \\ \left(\bar{\zeta}_1, \sum_{\iota=1}^{n} \bar{z}_{*\iota}, \bar{\xi}\right) &\in \Phi(\bar{p}, (\bar{\pi}_\iota)_{\iota=1}^{n}; e_N) \text{ for some } \bar{\zeta}_1 \Bigg\} \end{aligned}$$

for each $e_N \in \mathscr{E}_N$.

Now, a message correspondence $\nu_\theta: \mathscr{E} \to R^K$ is defined by

$$\nu_\theta(e) = \left(\bigcap_{l=1}^{n} \nu_\theta^l(e_l)\right) \cap \nu^N(e_N)$$

for each $e \in \mathscr{E}$. Furthermore, let $L_\theta = \nu_\theta(\mathscr{E})$ and define an outcome function $g_\theta: L_\theta \to Z$ by

$$g_\theta(\bar{m}) = (((\bar{z}_l, \bar{x}_l))_{l=1}^{n}, (\bar{\zeta}, \bar{\xi}))$$

for each $\bar{m} = (\bar{p}, (\bar{\pi}_l)_{l=1}^{n}, (\bar{z}_{*l})_{l=1}^{n}, \bar{\xi}, \bar{\sigma}) \in L_\theta$, where

$$\begin{aligned} \bar{z}_{1l} &= \theta_l \bar{\sigma} - \bar{p} \cdot \bar{z}_{*l} - \bar{\pi}_l \cdot \bar{\xi}, \qquad l = 1,\dots, n, \\ \bar{x}_l &= \bar{\xi}, \qquad l = 1,\dots, n, \\ \bar{\zeta}_1 &= \bar{\sigma} - \bar{p} \cdot \left(\sum_{l=1}^{n} \bar{z}_{*l}\right) - \left(\sum_{l=1}^{n} \bar{\pi}_l\right) \cdot \bar{\xi}, \end{aligned}$$

and

$$\bar{\zeta}_* = \sum_{l=1}^{n} \bar{z}_{*l}.$$

The triplet $\Omega_\theta = (L_\theta, \nu_\theta, g_\theta)$ will be called the *Lindahl process* on $\mathscr{E}$ with fixed profit shares $\theta = (\theta_i)_{i=1}^n$. The message correspondence ν_θ, in conjunction with the outcome function g_θ, yields a Lindahl equilibrium $(\bar{p}, (\bar{\pi}_i)_{i=1}^n, ((\bar{z}_i, \bar{x}_i))_{i=1}^n, (\bar{\zeta}, \bar{\xi}))$ for an environment $e \in \mathscr{E}$, i.e.,

(i) the joint action $(((\bar{z}_i, \bar{x}_i))_{i=1}^n, (\bar{\zeta}, \bar{\xi}))$ is consistent, i.e., it belongs to Z,

(ii) for each consumer $(i = 1, \ldots, n)$, $(\bar{z}_i, \bar{x}_i)$ is an element of $\Psi_i(\bar{p}, (\bar{\pi}_\iota)_{\iota=1}^n, \bar{\sigma}; e_i)$,

(iii) for the producer, $(\bar{\zeta}, \bar{\xi})$ is an element of $\Phi(\bar{p}, (\bar{\pi}_\iota)_{\iota=1}^n; e_N)$ and $\bar{\sigma} = \bar{\zeta}_1 + \bar{p} \cdot \bar{\zeta}_* + (\sum_{\iota=1}^n \bar{\pi}_\iota) \cdot \bar{\xi}$, for the environment $e \in \mathscr{E}$.

Now, the next lemma can be proved.

LEMMA 4. *Let $\Omega_\theta^E = (L_\theta^E, \nu_\theta^E, g_\theta^E)$ be the Lindahl process on E with fixed profit shares $\theta = (\theta_i)_{i=1}^n$ with $\theta_i > 0$ for all i $(i = 1, \ldots, n)$ and $\sum_{i=1}^n \theta = 1$. Then, Ω_θ^E is a privacy preserving and non-wasteful resource allocation process on E with respect to $\mathfrak{P}$. Furthermore, the message correspondence ν_θ^E is a homeomorphism from E onto the message space L_θ^E, and hence the message space L_θ^E is homeomorphic to R^K, where $K = N(s-1)+1$.*

Proof. Define an open subset D of R^K by

$$D = \Big\{ (p, (\pi_i)_{i=1}^n, (z_{*i})_{i=1}^n, \xi, \sigma) \in R_{++}^K \mid p \in R_{++}^{l-1}, \pi_i \in R_{++}^q \text{ and}$$

$$z_{*i} \in R_{++}^{l-1} \text{ for } i = 1, \ldots, n, \xi \in R_{++}^q, \sigma \in R_{++},$$

$$\theta_i \sigma - p \cdot z_{*i} - \pi_i \cdot \xi > 0 \text{ for } i = 1, \ldots, n, \text{ and}$$

$$\sigma - \sum_{j=2}^{l} p_j > 0 \Big\}.$$

It will be shown that $D = L_\theta^E$ and that ν_θ^E is a homeomorphism from E onto D. For a given environment $e = (((R_+^s, \rho(B^i), 0))_{i=1}^n, (Y^C, w_N)) \in E$, the only element $(\bar{p}, (\bar{\pi}_i)_{i=1}^n, (\bar{z}_{*i})_{i=1}^n, \bar{\xi}, \bar{\sigma})$ of $\nu_\theta^E(e)$ is given as:

$$p_j = c_j, \qquad j = 2, \ldots, l, \tag{23}$$

$$\bar{\pi}_{ki} = \frac{\beta_k^i \theta_i}{A_i} \Big/ \sum_{\iota=1}^{n} \frac{\beta_k^\iota \theta_\iota}{\gamma_k A_\iota} > 0, \qquad k = 1, \ldots, q; \quad i = 1, \ldots, n, \tag{24}$$

$$\bar{z}_{ji} = \theta_i \left(\omega + \sum_{\tau=2}^{l} c_\tau \right) \cdot \frac{b_j^i}{c_j A_i} > 0, \qquad j = 2, \ldots, l; \quad i = 1, \ldots, n, \tag{25}$$

$$\bar{\xi}_k = \sum_{i=1}^{n} \theta_i \left(\omega + \sum_{j=2}^{l} c_j \right) \cdot \frac{\beta_k^i}{\gamma_k A_i} > 0, \qquad k = 1, \ldots, q, \tag{26}$$

and

$$\bar{\sigma} = \omega + \sum_{j=2}^{l} c_j > 0, \tag{27}$$

where $A_i = 1 + \sum_{j=2}^{l} b_j^i + \sum_{k=1}^{q} \beta_k^i$ for $i = 1,\ldots, n$, and parameters are $B^i = (b_2^i,\ldots, b_l^i; \beta_1^i,\ldots, \beta_q^i) \in R_{++}^{s-1}$ $(i = 1,\ldots, n)$, $C = (c_2,\ldots, c_l; \gamma_1,\ldots, \gamma_q) \in R_{++}^{s-1}$, and $w_N = (\omega, 1,\ldots, 1)$. Now, it can be easily verified that $\theta_i\bar{\sigma} - \bar{p} \cdot \bar{z}_{*i} - \bar{\pi}_i \cdot \bar{\xi} = \theta_i(\omega + \sum_{j=2}^{l} c_j)/A_i > 0$ for $i = 1,\ldots, n$, and that $\bar{\sigma} - \sum_{j=2}^{l} \bar{p}_j = \omega > 0$. Consequently, the equilibrium message $(\bar{p}, (\bar{\pi}_i)_{i=1}^n, (\bar{z}_{*i})_{i=1}^n, \bar{\xi}, \bar{\sigma})$ belongs to the set D defined above. The expressions (23)–(27) imply that the message correspondence ν_θ^E is a continuous mapping from E into D.

Conversely, take an arbitrary element $(p, (\pi_i)_{i=1}^n, (z_{*i})_{i=1}^n, \xi, \sigma)$ of D. Then, the only element $e = (((R_+^s, \rho(B^i), 0))_{i=1}^n, (Y^C, w_N))$ of the inverse $\nu_\theta^{E\,-1}(p, (\pi_i)_{i=1}^n, (z_{*i})_{i=1}^n, \xi, \sigma)$ can be given by parameters:

$$b_j^i = p_j z_{ji}/(\theta_i\sigma - p \cdot z_{*i} - \pi_i \cdot \xi) > 0, \qquad j = 2,\ldots, l; \quad i = 1,\ldots, n, \tag{28}$$

$$\beta_k^i = \pi_{ik}\xi_k/(\theta_i\sigma - p \cdot z_{*i} - \pi_i \cdot \xi) > 0, \qquad k = 1,\ldots, q; \quad i = 1,\ldots, n, \tag{29}$$

$$c_j = p_j > 0, \qquad j = 2,\ldots, l, \tag{30}$$

$$\gamma_k = \sum_{i=1}^{n} \pi_{ki} > 0, \qquad k = 1,\ldots, q, \tag{31}$$

and

$$\omega = \sigma - \sum_{j=2}^{l} p_j > 0. \tag{32}$$

This implies that the inverse $\nu_\theta^{E\,-1}$ is a continuous mapping from D into E. Consequently, the mapping ν_θ^E is a homeomorphism from E to $D = L_\theta^E$. This fact shows incidentally that the Lindahl process Ω_θ^E is a resource allocation process on E. Evidently, it is privacy preserving and non-wasteful on E with respect to $\mathfrak{P}$. Q.E.D.

It will be clear that, if L_θ is the Lindahl message space for a topological space $\mathscr{E}$ of environments, then $R^K \geqslant_F L_\theta$, since L_θ is a subspace of R^K. This observation and Theorem 1 (and Theorem 2) imply the following theorem.

THEOREM 3. *Let $\mathscr{E}$ be a topological space of environments which includes $(E, \mathscr{T})$ as a subspace. Let $\Pi = (M, \mu, h)$ be a resource allocation process on $\mathscr{E}$ that is privacy preserving and non-wasteful with respect to $\mathfrak{P}$*

on E, and let $\Omega_\theta = (L_\theta, \nu_\theta, g_\theta)$ *be the Lindahl process on* $\mathscr{E}$*. Suppose that the process* Π *satisfies either of the following three conditions:* (1) *M is a Hausdorff space;* (2) *the inverse of the restriction* $\mu|_E$ *is upper hemi-continuous*; (3) *the inverse of* $\mu|_E$ *is lower hemi-continuous. Then, it holds that* $M \geqslant_F L_\theta$.

It should be noted that the Lindahl process is actually a privacy preserving and non-wasteful resource allocation process on a space which includes E as a subspace.[15] Then, Theorem 3 states that the size of the Lindahl message space is the minimal informational size of message spaces for non-wasteful and privacy preserving resource allocation processes on a space of environments which includes E as a subspace.

Remark. In the definition of the Lindahl process, we supposed that each consumer determined proposed production of public goods taking prices as given. Another formulation of the behavior of consumers might be possible, where each consumer reports the (marginal) valuation for the public goods meeting proposed production of the public goods. When only *equilibrium* messages matter as in this paper, the two formulations make no differences.[16] Out of equilibrium, however, informational burden of the process is different depending on formulations. A related question was pointed out by Aoki [1].

5. CONCLUDING REMARKS

In this paper, it has been shown that the minimal informational size of message spaces for privacy preserving and non-wasteful resource allocation processes on a certain space of environments with public goods is the size of the Lindahl message space. This is a kind of impossibility theorem; i.e., there exists no privacy preserving and non-wasteful resource allocation process over a certain space of environments that uses message space whose informational size is smaller than that of the Lindahl message space.

Theorem 1 of this paper might be considered to be an extension of the modified version of Theorem 35 of Mount and Reiter [14] (Theorem 1 of Walker [17]), while Theorem 2 of this paper is an extension of the theorem of Osana [15], to the case of economies containing public goods.

Certainly, as a special case, the method of this paper can be applied to the case of economies containing only private goods only with a slight modification.[17] Then, it could be shown that the minimal informational size is that of $R^{N(l-1)+1}$, and that this size is that of the competitive message space. It is already known that in pure exchange economies with N agents and l commodities the minimal informational size is that of $R^{N(l-1)}$. (See [15].) The plus one here comes from an additional message as to the profit of the producer in the case of the competitive process.

---

[15] For instance, let $\mathscr{E}$ be the space where the Lindahl message correspondence is upper hemi-continuous and is a mapping. Then, the message correspondence is locally threaded (actually continuous).

[16] Note that, at the Lindahl equilibrium defined above, equal quantities of the services of the public goods are consumed by all the consumers.

[17] For instance one can take $w_N = (\omega, 0, \dots, 0)$ in the definition of the subclass E $(\omega > 0)$. The proof is the same as in the text.

In this paper, results were obtained for the case of a single producer. In the case of several producers, the results cannot be derived by simply utilizing the Cobb–Douglas=Linear (CDL) economies as a subspace E. The difficulty is that there are immense possibilities of "corner solutions" when there are several producers with linear production funtions (or, in other words, when there are several producers whose production possibility sets are cones with vertices $(w_f, 0)$). And the injectiveness lemma (Lemma 3), which is the key lemma, fails to hold in this case.

This difficulty could be avoided by modifying the definition of the subspace E. To see this, suppose that there are m producers, l private goods, and one public good ($q = 1$), and that the private goods are used as inputs to produce the public good which is the only output. In order to define the new subclass E, the class of characteristics of the consumer is defined similarly to the case of CDL economies, and the class of characteristics of the producer is defined by the product of the production possibility sets represented by the Cobb–Douglas production functions and the set of vectors of endowments of private goods whose first component varies over the positive reals and other components are fixed (to unity for example). The space E thus defined and identified with the product space of the parameter spaces of Cobb–Douglas functions is homeomorphic to $R^{N(s-1)+m}$ (where $N = m + n$ and $s = l + 1$). And the injectiveness lemma (Lemma 3) holds again for this space E because "corner solutions" do not occur owing to the fact that the marginal *productivities of the Cobb–Douglas production functions tend to plus* infinity at the origin. It is clear that the informational size of $R^{N(s-1)+m}$ is that of the Lindahl message space. The plus m corresponds to the number of the messages about the profits of the producers. These results would easily be generalized to the case of an arbitrary number of public goods.

ACKNOWLEDGMENTS

I would like to thank Professors H. Haga, M. Otsuki, and H. Hori of Tohoku University for their stimulation given to me. I am also very grateful to the anonymous referees for their constructive and helpful comments. The responsibility for any error rests with me.

REFERENCES

1. M. AOKI, Two planning processes for an economy with production externalities, *Internat. Econ. Rev.* **12** (1971), 403–414.

2. X. Calsamiglia, Decentralized resource allocation and increasing returns, *J. Econ. Theory* **14** (1977), 263–283.
3. T. Groves and J. Ledyard, Optimal allocation of public goods: a solution to the "free rider" problem, *Econometrica* **45** (1977), 783–809.
4. F. A. Hayek (Ed.), "Collectivist Economic Planning," Routledge and Kegan Paul, London, 1935.
5. W. Hildenbrand, On economies with many agents, *J. Econ. Theory* **2** (1970), 161–188.
6. W. Hildenbrand, "Core and Equilibria of a Large Economy," Princetion Univ. Press, Princeton, 1974.
7. L. Hurwicz, Optimality and informational efficiency in resource allocation processes, *in* "Mathematical Methods in Social Sciences" (K. J. Arrow, S. Karlin and P. Suppes, Eds.), 27–46, Stanford Univ. Press, Stanford, 1960.
8. L. Hurwicz, On informationally decentralized systems, *in* "Decision and Organization" (C. B. McGuire and R. Radner, Eds.), North-Holland, Amsterdam, 1972.
9. L. Hurwicz, On the dimensional requirements of informationally decentralized Pareto-satisfactory processes, *in* "Studies in Resource Allocation Processes" (K. J. Arrow and L. Hurwicz, Eds.), Cambridge Univ. Press, Cambridge, 1977.
10. J. L. Kelley, "General Topology," Van Nostrand, Princeton, N. J., 1955.
11. T. C. Koopmans, "Three Essays on the State of Economic Science," McGraw–Hill, New York, 1957.
12. O. Lange, On the economic theory of socialism, *in* "On the Economic Theory of Socialism" (B. E. Lippincott, Ed.), Univ. of Minnesota Press, Minneapolis, 1938.
13. J. C. Milleron, Theory of value with public goods: a survey article, *J. Econ. Theory* **5** (1972), 419–477.
14. K. Mount and S. Reiter, The informational size of message spaces, *J. Econ. Theory* **8** (1974), 161–192.
15. H. Osana, On the informational size of message spaces for resource allocation processes, *J. Econ. Theory* **17** (1978), 66–78.
16. F. Sato, A note on the topology of closed convergence for a class of preferences represented by utility functions with parameter, unpublished paper, 1979.
17. M. Walker, On the informational size of message spaces, *J. Econ. Theory* **15** (1977), 366–375.

JOURNAL OF ECONOMIC THEORY 24, 70–94 (1981)

On the Design of Regulatory Price Adjustment Mechanisms*

DAVID P. BARON

Northwestern University, Evanston, Illinois 60201

AND

RAYMOND R. DE BONDT

Katholieke Universiteit Leuven, Leuven, Belgium

Received March 1, 1979; revised April 7, 1980

Price adjustment mechanisms are employed in the electric utility industry to pass changes in fuel costs on to consumers without formal rate review by a regulatory commission. The predictability of this pass-through and the regulator's limited ability to observe the actions of a firm can create potential incentive problems associated with the choices of technology and fuel supply. This paper is concerned with the regulatory design of pass-through formulas when a factor price is uncertain. The optimal design involves deviating from the full-information optimal price formulas in order to mitigate the incentive problems.

1. INTRODUCTION

During the 1950's and 1960's the process for electric utility regulation functioned smoothly as firms were able to earn a satisfactory return while not increasing, and often decreasing, the nominal prices for their output. Increases in factor prices were offset by productivity gains and by the effects of economies of scale realized through growth in demand and through elasticity responses to decreases in real prices. The acceleration of factor price increases that began in the late 1960's, and peaked after the oil price increases in 1973, however, necessitated changes in output prices, and regulatory commissions were forced to deal with a rapidly increasing case load. According to Joskow (1974), "Rapid inflation had quickly changed a very passive and inactive 'rate of return' regulatory process into a very active and continual process of administrative rate of return review. The regulatory

---

* This research has been supported by the National Science Foundation, Grant SOC 77-07251.

0022-0531/81/010070-25$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved

process, in terms of both its techniques for adjustments and the staff inputs available to implement the techniques, was completely unsuited for the new economic environment (p. 314)." One feature of electric utility tariffs that expedited rate changes during this period is a fuel adjustment clause (FAC) that permits changes in the average cost of fuels to be passed on directly to consumers without specific commission review. As an indication of the extent to which FAC's are used, NARUC (1977, p. 315) reported that "An FPC survey as of January 1, 1974, showed that 65 % of the larger privately-owned utilities had fuel adjustment clauses in their residential schedules, 77 % had such clauses in their commercial schedules, and 83 % in their industrial schedules."[1] The total documented increases in rates attributable to FAC's for the commissions responding to the Congressional Research Service survey (U.S. Senate, 1977) for the years from 1973 to 1976 were $1.0 billion, $2.6 billion, $6.4 billion, and $5.2 billion, respectively.

Proponents of fuel adjustment clauses argue that these mechanisms are needed to moderate the deterioration in the rate of return during the "processing lag" period associated with an administrative rate review. Supporters also argue that the use of FAC's provides a better matching between revenues and costs than can be achieved with infrequently adjusted prices. FAC's may also yield benefits by reducing the case load so that commissions can deal with their other regulatory responsibilities such as prior certification of investments and approval of financing programs.

The widespread use of FAC's and the magnitude of the resulting price increases have, however, led to strong criticism of these adjustment clauses. The House Subcommittee on Oversight and Investigations (1975, p. 1) concluded:

> ... the Subcommittee finds that fuel adjustment clauses, in any form, are unwise, unnecessary, unworkable, and unfair. Fuel clauses, in the Subcommittee's view, moreover, may be the forerunner of a campaign to undermine utility regulation by means of a full range of automatic adjustment clauses that will render utility commissions virtually powerless while reducing utility management to the not so fine art of adjusting monthly consumer electric bills to reflect the industry's unchecked cost increases.

In addition to this abdication of regulatory authority criticism, use of an adjustment formula may engender incentive problems if there are profit gains that can result from endogenous cost increases. These incentive problems may also be present under a system of administrative rate review, accompanied by interim rate increases when necessary, but the uncertainty regarding the amount of a rate increase that might be granted in response to a cost increase could dampen this incentive. In addition to the shorter time required for a rate adjustment, this predictability of the price adjustments resulting from a fuel adjustment clause distinguishes it from the administrative rate review process.

---

[1] In 1971 FAC's were allowed in all but five states, Montana, Oregon, Utah, Washington, and Wyoming, and in Idaho and Nevada "FAC's are not used or denied to some categories of utilities (U. S. Senate (1977), p. VI)." Of the 82 electric utilities responding to an October 3, 1974 questionnaire of the Edison Electric Institute (1974), 81 reported that a FAC was included in their tariffs.

To illustrate the nature of the incentive problem, consider the type of "average cost" fuel adjustment clause currently being utilized. If p_0 denotes the current price and F_0 the quantity of fuel utilized at an output level $Q(p_0)$, the price p adjusted in response to a factor price change from c_0 to c is given by the formula

$$p = p_0 + (c - c_0)F_0/Q(p_0). \tag{1}$$

If the profit of the firm is increasing in price at p, a potential moral hazard problem is present to the extent that the firm can affect the fuel usage F_0, and hence the price adjustment, through its choice of technology. A second potential incentive problem centers on the firm's ability to affect the price adjustment through the factor price c of the fuel purchased. A firm may not have an incentive to purchase the least costly fuels, or to convert from one type of fuel to another, if the profit effect of the resulting output price is greater than the additional fuel cost. For example, the New Mexico Public Service Commission (1975, p. 119) has stated

> Under such authorizations, the electric utility can pass on its costs of fuel and purchased power to its customers more or less contemporaneously with experienced increases. However, since it cannot correspondingly increase its service rates to cover increases in depreciation and capital costs per unit of generating capacity or per unit of energy sold, it is disinclined to shift from gas and oil fired generation to coal and nuclear systems. In short, while automatic cost of fuel and purchased power adjustment authorizations are needed to enable the electric utility "to meet all costs of furnishing services," they are likely to operate as a positive disincentive to prepare for and improve the future.

The purpose of this study is to investigate the optimal design of automatic adjustment clauses in light of the potential incentive problems they entail.[2] If the regulator could perfectly observe or monitor the technology and fuel supply decisions of the firm, the regulator could in principle use its certification authority and other statutory powers to force the firm to make the technically efficient choices. In practice, regulatory commissions are not equipped to perform this monitoring but instead observe accounting and engineering data and use their certification authority in a more restrictive manner than to dictate technology choices. Furthermore, a regulator is unlikely to be able to observe the range of fuel factor prices that a firm faces and may not be able to observe the firm's fuel supply choices without extensive auditing or costly reporting by the firm. Instead, regulatory commissions appear to rely on accounting data that record the "incurred" cost of those fuels actually purchased. When observability is limited, the regulator is faced with a second-best design problem which is considered in Sections 4 and 5 after the model and the full-information solution have been presented.

---

[2] A number of practical aspects of the design problem such as dealing with line losses, heat rates, fuel mix, fuel-related costs, etc., will not be treated here. Furthermore, dynamic aspects of the design of automatic adjustment mechanisms, such as incorporating provisions for productivity changes, technological change, and demand growth, will not be considered.

2. THE MODEL

The design of adjustment formulas will be considered in the context of one-part and two-part pricing systems, since they may be utilized in the pricing of many utility services without necessitating special metering.[3] Individual consumers will be identified in terms of their type z, and the demand function for an individual of type z will be denoted $Q^+(p, z)$, when a unit price p is charged.[4] Consumers are assumed to be ordered by $z(z \in [0, \bar{z}])$ with increasing values of z corresponding to higher demand for every p, so that $Q^+(p, z) > Q^+(p, z')$ for all p when $z > z'$.[5] The distribution of consumers will be represented by a histogram $n(z)$, so that the number $N(z^0)$ of type z^0 or more is

$$N(z^0) \equiv \int_{z^0}^{\bar{z}} n(z)\, dz.$$

The distribution of consumer types will be assumed to be such that a consumer of type $z = 0$ will not purchase at any $p > 0$.[6]

---

[3] Two-part price structures have been considered by Hotelling (1938), Baumol and Bradford (1970), Diamond and Mirrlees (1971), Oi (1971), Ng and Weisser (1974), and Faulhaber and Panzar (1977), but they have not considered the choice of technology as will be treated here. These pricing structures may be extended by considering unit prices that depend on the quantity consumed as considered by Leland and Meyer (1976), Spence (1977), Faulhaber and Panzar, and Roberts (1979).

[4] The model of consumer demand is the same as that utilized by Leland and Meyer and Faulhaber and Panzar. The individual demand functions are specified as independent of an individual's income in order to avoid complications in the definition of consumer surplus. This assumption of zero income elasticity can also be thought of as an approximation justified by Willig's (1976) study of consumer surplus measures.

[5] This assumption implies that the demand functions for individuals of different types do not cross which eliminates the case of prices below marginal cost when returns to scale are increasing as considered by Oi and by Ng and Weisser.

[6] This assumption permits a simplification of the statement of the results.

For a two-part price structure with a fixed fee T a consumer will purchase if his surplus is at least as great as T or if

$$\int_{p}^{\infty} Q^{+}(p', z)\, dp' \geqslant T. \tag{2}$$

The marginal consumer will be of type $z^0 = z^0(p, T)$ defined by the equality in (2), and z^0 is a strictly increasing function of p and T, since direct calculation yields $\partial z^0/\partial p = Q^{+}(p, z^0)\, \partial z^0/\partial T > 0$. When a one-part price structure ($T = 0$) is utilized, the marginal consumer z^0 is the largest z such that $Q^{+}(p, z) = 0$. Aggregate demand $Q(p, z^0)$ for either price structure is thus given by

$$Q(p, z^0) \equiv \int_{z^0}^{\bar{z}} Q^{+}(p, z)\, n(z)\, dz,$$

and the number N of consumers purchasing from the firm, and hence paying the fixed fee, is $N(z^0) = N(z^0(p, T))$.

A single-period model will be utilized in which the firm makes its choice of a technology before, and its choice of a fuel supply after, the change in the factor price and the resulting output price adjustment. The choice of technology may represent the choice among generation technologies or for an existing generation plant may represent, for example, the investment in improved transmission equipment to reduce line losses. The technology will be assumed to involve two inputs, fuel f and capital k, and the cost function will be expressed as $\varphi(Q)\, M(c, \gamma)$, where $\gamma \equiv f/k$ is the fuel-capital ratio, $\varphi(\cdot)$ is a strictly increasing, continuously differentiable function expressing returns to scale effects, and M is the cost per unit of $\varphi(Q)$. The function M can be expressed as $M(c, \gamma) = cf(\gamma) + rk(\gamma)$, where r is the factor price of capital and $f(\gamma)$ and $k(\gamma)$ are the quantities of fuel and capital employed per unit of $\varphi(Q)$. The unit cost function M is assumed to be strictly increasing in c and γ, and strictly concave and twice differentiable in γ. As indicated in Baron and De Bondt (1979), a cost function of this form can be derived from a putty-clay technology with a homothetic ex ante production function and an ex post production function with the fuel and capital inputs employed in the fixed ratio γ.

The profit $\pi(p, T, c, \gamma)$ of the firm is

$$\pi(p, T, c, \gamma) \equiv TN(z^0) + pQ - \varphi(Q)\, M(c, \gamma), \tag{3}$$

and when a one-part price structure is utilized, profit is given by (3) with $T = 0$. The objective of the firm is assumed to be the maximization of expected profit $E\pi$ where c, and hence p and T, are uncertain at the time the fuel-capital ratio is chosen. In order to ensure that the firm can attract capital, the regulator must choose a price adjustment formula so that the providers of capital receive a fair return. Since uncertainty regarding the future factor price is present, an adjustment mechanism can be used to either provide investors with an ex ante or an ex post fair return. An ex ante fair return would yield investors an "expected" return commensurate with returns available elsewhere in the capital market, while an ex post fair return would require that the return be fair for any state of nature that may result. Since regulation in practice does not and should not guarantee an ex post return, ex ante regulation will be taken to characterize regulatory policy. The regulator will thus be constrained in its design of adjustment formulas by the requirement that expected profit be nonnegative or $E\pi \geqslant 0$.

The objective of the regulator can be considered from a descriptive or a prescriptive point of view. As a description of regulatory practice Joskow (1974) concluded that the regulation of electric utilities at the state level is characterized by the objective of holding down nominal prices or the rate of nominal price increases. In the context of the model developed here this policy could be represented by the objective of setting the lowest possible price p subject to the restriction that the owners of the firm are allowed to earn a fair return on their capital. This objective is equivalent to the maximization of consumer surplus, but the objective of maximizing consumer plus producer surplus is more appropriate from a normative point of view. As Bailey (1976, p. 394), however, states in rejecting the consumer plus producer surplus objective:

> To suggest to a commission that it might wish to maximize an objective including the benefits to producers is to suggest a sort of behavior that smacks of Stigler's 'capture' theory of regulaltory agencies, and is at odds with the 'public-interest' view with which a commission is likely to pride itself [see Posner (1974)]. Thus, my reason for choosing the objective that considered only the position of consumers is that it seemed closer to the stated charter of a regulatory commission.

Since the purpose here is to study adjustment mechanisms in the context of the existing regulatory environment, the regulator will be assumed to have consumer surplus objectives, where the surplus $S(p, T)$ is given by $S(p, T) \equiv \int_p^\infty Q(p', z^0)\, dp' - TN(z^0)$. The prices to be paid by consumers under an automatic adjustment mechanism are uncertain, however, so expected consumer surplus

$$ES = \int S(p, T)\, h(c)\, dc$$

will be utilized.[7] Since the emphasis here is on the design of automatic adjustment mechanisms and on the resulting incentive problems caused by imperfect observability, the regulator and the firm will be assumed to have ex ante asymmetric information regarding the future factor price. That information will be represented by a (differentiable) density function $h(c)$ of the factor price. In the next two sections the fuel supply is assumed to be exogenously determined and the choice of technology is analyzed, while in Section 5 the fuel supply decision itself is treated.

---

[7] This measure is used as an approximation. The limitations of this measure are considered by Schmalensee (1972).

3. THE FULL-INFORMATION SOLUTION AND THE CHOICE OF TECHNOLOGY

The full-information solution is that which the regulator would select given complete information regarding the firm, its choices, and the factor prices it faces. When the regulator has complete information, he can utilize a "forcing mechanism" of the form

$$\begin{aligned}(\bar{p}(c,\gamma), \bar{T}(c,\gamma)) &= (p(c), T(c)) \qquad \text{if} \quad \gamma = \gamma^* \\ &= (0, 0) \qquad\qquad\quad \text{if} \quad \gamma \neq \gamma^*,\end{aligned}$$

which forces the firm will choose γ^* and to operate with price functions $p(c)$ and $T(c)$. Conditionally on γ^* the regulator will choose the price functions according to the program

$$\max_{p(c),T(c)} \int S(p(c), T(c))\, h(c)\, dc$$

$$\text{subject to} \quad \int \pi(p(c), T(c), c, \gamma^*)\, h(c)\, dc \geqslant 0.$$

This program may be reformulated by considering the maximization of consumer surplus for a constrained level of ex post profit ρ and then choosing the optimal profit function $\rho(c)$. Define the optimized ex post consumer surplus $S^*(\rho, c, \gamma)$ for a given γ^* by

$$S^*(\rho, c, \gamma^*) \equiv \max_{p,T} \{S(p, T) \text{ subject to } \pi(p, T, c, \gamma^*) = \rho\}. \tag{4}$$

The profit function $\rho(c)$ is then determined according to the program

$$\begin{aligned}&\max_{\rho(c)} \int S^*(\rho(c), c, \gamma^*)\, h(c)\, dc \\ &\text{subject to} \quad \int \rho(c)\, h(c)\, dc \geqslant 0.\end{aligned} \tag{5}$$

Pointwise optimization yields the first-order condition[8]

$$S_\rho^* + \lambda^* = 0, \tag{6}$$

where $\lambda^* > 0$ is the Lagrangian multiplier associated with the expected profit constraint in the program in (5). From the solution of (4)

$$S_p = S_\rho^* \pi_p,$$
$$S_T = S_\rho^* \pi_T \qquad \text{if} \quad T(c) > 0,$$

so the optimality conditions for $\rho(c)$ and $T(c)$ may be stated as[9]

$$-\frac{S_p}{\pi_p} = \frac{Q}{Q + (p - \varphi' M)\, Q_1 - \pi_{z0} n(z^0)(\partial z^0/\partial p)} = \lambda^* \tag{7}$$

$$-\frac{S_T}{\pi_T} = \frac{N(z^0)}{N(z^0) - \pi_{z0} n(z^0)(\partial z^0/\partial T)} = \lambda^* \qquad \text{if} \quad T > 0, \tag{8}$$

where $\pi_{z0} = T + (p - \varphi' M)\, Q^+$, $Q^+ = Q^+(p, z^0)$, $Q_1 = \partial Q/\partial p$, and $\varphi' = \varphi'(Q^+)$. Optimality of the functions $\rho^*(c)$, $p^*(c)$, and $T^*(c)$ characterized by (6)–(8) requires that $S + \lambda\pi$ be concave at (p^*, T^*) for all c. The profit function is at least locally optimal if S^* concave in ρ at ρ^* for all c, which requires that

$$S_{\rho\rho}^* = \frac{\partial}{\partial \rho}\left(\frac{S_p}{\pi_p}\right) < 0 \ \text{ at } \ \rho(c) = \rho^*(c), \qquad \text{for all } c. \tag{9}$$

Consequently, if the ratio of marginal consumer surplus to marginal profit is decreasing with ex post profit, ρ^* yields a local optimum. These conditions will be assumed to be satisfied.

The conditions in (6)–(8) characterize the Pareto optimal sharing of the risk resulting from the uncertain factor price and indicate that it is optimal for both parties to bear a portion of that risk. Since an increase in expected profit reduces expected consumer surplus, $\lambda^* > 0$ and (5) indicates that at ρ^* an increase in expected profit reduces expected consumer surplus. This implies from (7) that with the optimal unit price function p^*, marginal profit π_p is positive, so regulation is effective in maintaining price below the level that the firm would set given γ^*. A similar result follows from (8) if $T^* > 0$. These conclusions are stated as

---

[8] If the regulatory objective is the maximization of total surplus, the multiplier λ in (6) is replaced by $(\lambda + 1)$ and an analogous change is made in the conditions characterizing the second-best solution in Section 4.

[9] The derivative of consumer surplus with respect to T is $\partial S/\partial T = -N(z^0) + (\int_p^\infty Q^+(p', z^0)\, dp' - T)\, n(z^0)(\partial z^0/\partial T)$. By definition of z^0, the term $(\int_p^\infty Q^+(p', z^0)\, dp' - T)$ equals zero.

Proposition 1. *The full-information adjustment formulas are such that regulation is effective in restricting the unit price $p^*(c)$ and the fixed fee $T^*(c)$ (when $T^*(c) > 0$) below the levels that the firm prefers given the fuel-capital ratio γ^*, i.e., $\pi_p > 0$, and $\pi_T > 0$ if $T^*(c) > 0$.*

As is indicated in the following proposition, a two-part price structure is strictly preferred to a one-part price structure only when the technology is characterized by increasing returns to scale. The proof is provided in Appendix A, and a further characterization of the optimal price structure is given in Appendix B.

Proposition 2. *If returns to scale are increasing (decreasing) (constant) for all $Q(p^*, z^0)$, then $\lambda^* > (<)(=) 1$, $T^*(c) > (=)(=) 0$, and $p^*(c) > (<)(=) \varphi' M$, where $\varphi' M$ is marginal cost.*

A two-part tariff will thus only be utilized for firms with increasing returns to scale. To characterize the full-information price function in this case, substitute $\partial z^0/\partial T = (\partial z^0/\partial p)/Q^+$ into (8) and combine with (7) to obtain

$$-(1 - \lambda^*)(Q - N(z^0)\, Q^+) + \lambda^*(p^* - \varphi' M)\, Q_1 = 0. \tag{10}$$

Then, divide by Q_1, let $\varepsilon = pQ_1/Q < 0$ denote the price elasticity of demand, let $s(z^0) \equiv Q^+/Q$ be the share of aggregate demand represented by the marginal consumer, and rewrite to obtain

$$p^* \left(1 + \frac{(1 - 1/\lambda^*)}{\varepsilon} (1 - N(z^0)\, s(z^0))\right) = \varphi' M. \tag{11}$$

This expression is analogous to Oi's (1971) Eq. (10) except for the term $(1 - 1/\lambda^*)$ which replaces the constant one in his result. Since $\lambda^* > 1$ and $N(z^0)\, s(z^0) < 1$, p is greater than marginal cost for all c. Because $T^*(c) > 0$ and $p^*(c) > \varphi' M$, the term $\pi_{z^0} = (T^*(c) + (p^*(c) - \varphi' M)\, Q^+)$, which is the contribution to profit of the marginal consumer, is positive.

The full-information price functions depend on the factor price c, and the following proposition establishes that when a one-part price structure is utilized, a higher factor price results in a higher output price.

Proposition 3. *The full-information, one-part price function $p^*(c)$ is a strictly increasing function of the factor price c.*

Proof. Differentiation of (10) with $Q^+ = 0$ yields

$$\frac{dp^*}{dc} = \lambda^* \varphi' Q_1 \frac{\partial M}{\partial c} \bigg/ B, \tag{12}$$

where

$$B = -(1 - 2\lambda^*)\,Q_1 - \lambda(\varphi'' Q_1^2 M + (p^* - \varphi' M)\,Q_{11})$$

is the second-order condition which has been assumed to be negative and $Q_{11} = \partial^2 Q/\partial p^2$. The numerator is negative, so $dp^*/dc > 0$. ▮

Although the full-information one-part price function is an increasing function of the factor price, it is not possible to determine in general how the full-information, two-part price functions are related to c for the case of increasing returns to scale.

The form of the full-information profit function $\rho^*(c)$ is important for the fuel supply decision considered in Section 5 and can be determined by multiplying (11) by Q, adding T^*N and subtracting φM on both sides, and rearranging to obtain

$$\rho^*(c) = \varphi M(1/\eta - 1) + (1/\lambda^* - 1)\,Q^2(1 - N(z^0)\,s(z^0))/Q_1 + T^*N(z^0), \qquad (13)$$

where $\eta \equiv \varphi/(\varphi' Q)$ is the measure of returns to scale. The form of $\rho^*(c)$ is difficult to determine in general, but for constant returns to scale ($\eta = 1$) Proposition 2 implies that $\rho^*(c) = 0$ for all c. With a one-part price structure, the derivative of $\rho^*(c)$ is

$$\rho^{*\prime}(c) = \varphi'' Q_1 Q p^{*\prime}(c) M + \varphi\left(\frac{1}{\eta} - 1\right)\frac{\partial M}{\partial c}$$
$$+ (1/\lambda^* - 1)((2Q_1^2 Q - Q_{11}Q^2)/Q_1^2)\,p^{*\prime}(c).$$

With increasing (decreasing) returns to scale the first term is positive (negative) and the second term is negative (positive). If $(2Q - Q_{11}(Q/Q_1)^2) > 0$, the third term is positive (negative). In general, the sign of $\rho^{*\prime}(c)$ is ambiguous.

As a particularly simple example, consider the case in which the demand function has constant elasticity ε ($\varepsilon < -1$) and a homogeneous production function, so that $\varphi(Q) = Q^{1/\eta}$. Substituting $\varepsilon = Q_1 p^*/Q$ into (13), taking the expectation, and equating expected profit to zero yields

$$(1/\eta - 1)\,E(\varphi M) + (1/\lambda^* - 1)\,E(pQ)/\varepsilon = 0. \qquad (14)$$

Since $E(\varphi M) = E(p^* Q)$ from the expected profit constraint, (14) implies that

$$1 - 1/\eta = (1/\lambda^* - 1)/\varepsilon.$$

Substituting this condition into (7) implies that

$$p^*(c) = M^{\lambda^*},$$

where $\lambda^* > (=)(<) 1$ if returns to scale are increasing (constant) (decreasing) (i.e., $\eta > (=)(<) 1$) as shown in Appendix B. Consequently, the price function is a strictly convex (concave) function of M if the firm has a technology exhibiting increasing (decreasing) returns to scale. The profit function in either case can then be shown to be identically zero for all M.

4. THE DESIGN PROBLEM AND OBSERVABILITY

The full-information price formulas are completely specified by evaluating (7) and (8) at the optimal fuel-capital ratio determined by maximizing with respect to γ in the program in (5) evaluated at $p^*(c)$. This is equivalent to minimizing expected cost $E[\varphi(Q(p^*(c), z^0)) M(c, \gamma)]$. Letting γ^* denoted the first-best fuel-capital ratio, the first-best price for the case of constant returns to scale, for example, is the linear function $p^*(c) = M(c, \gamma^*) = cf(\gamma^*) + rk(\gamma^*)$.

The full-information adjustment formulas can be implemented if the factor price c and the firm's choice of technology γ can be observed ex post. Since there may be costs associated with monitoring γ, it is natural to ask if there are any potential gains to that monitoring when the factor price itself can be observed. Given any price functions specified by the regulator, the firm will maximize its expected profit, but if the price functions depend only on c, it is clear the firm will choose the γ that minimizes its expected cost, and hence given $p^*(c)$ and $T^*(c)$ will choose γ^*. Consequently, when the factor price can be observed, there are no gains to monitoring the technology decisions of the firm. The observation of the factor price itself may not be possible, however, without significant expenditures by the regulator. For example, to observe c the regulator may have to audit the accounts of the firm to determine the factor price that actually occurred. Furthermore, if the source of information for a regulator is the accounting records of the firm, data are only available on the fuel supply sources actually chosen and not on those supplies that were available to but not purchased by the firm.

Corresponding to this case, the regulator will be assumed to be able to observe ex post only the "unit" cost $M(c, \gamma)$ and thus must utilize functions $p(M)$ and $T(M)$. This case is of particular interest because the adjustment formula in (1) can be shown to be of this form. Consider a one-part price structure of the form $p(M) = a + bM$. If $M_0 = M(c_0, \gamma)$ represents the initial unit cost before the factor price change, let $a = -bM_0 + p_0$ and $b = \varphi(Q(p_0))/Q(p_0)$, where p_0 is the initial output price corresponding to M_0. The adjusted price $p(M)$ is then

$$p(M) = p_0 + (\varphi(Q(p_0))/Q(p_0))(M - M_0) = p_0 + (c - c_0)F_0/Q(p_0),$$

where $F_0 = \varphi(Q(p_0))f(\gamma)$, which is the adjustment formula in (1).

An alternative interpretation of this average cost adjustment mechanism is as an adjustment triggered by a failure of the firm to earn its allowed return. If the firm has a price p_0 in effect and the factor price changes from c_0 to c, the profit π^+ of the firm prior to the output price adjustment is $\pi^+ = p_0 Q(p_0) - \varphi(Q(p_0))\, M(c, \gamma)$, and the deficit per unit of output is $\pi^+/Q(p_0)$. If the adjustment Δp in price covers this per unit deficit, the adjusted price p is

$$p = p_0 + \Delta p = p_0 - \pi^+/Q(p_0).$$

If p_0 was originally set equal to average cost $(p_0 = (\varphi(Q(p_0))/Q(p_0))\, M_0$, this expression reduces to (1).

Since the argument M of the adjustment formulas depends on the technology selected by the firm, a potential incentive problem is present. To analyze the optimal design of $p(M)$ and $T(M)$ in light of this potential incentive problem, it is convenient to transform the original problem in (4) and (5) from one involving a distribution of c to one involving a distribution of M.

Letting $y(M; \gamma)$ denote the density of unit cost given γ, expected profit

$$E\pi = \int (T(M)\, N(z^0) + p(M)\, Q(p(M))$$

$$- \varphi(Q(p(M)))M)\, y(M; \gamma)\, dM = \int \pi y\, dM$$

depends on γ only through the density function y.[10] In designing the price functions, the regulator must take into account the firm's choice of γ, so the regulator will maximize expected consumer surplus subject to the constraints that expected profit is nonnegative and that the firm maximizes its expected profit. The regulator's program is equivalent to choosing $\rho(M)$ according to

$$\max_{\rho(M)} \int S^*(\rho(M), M, \gamma)\, y\, dM, \tag{15}$$

$$\text{subject to} \int \rho(M)\, y\, dM \geqslant 0, \tag{16}$$

$$\gamma \in \underset{\gamma \in [0,\infty)}{\operatorname{argmax}} \int \rho(M)\, y\, dM, \tag{17}$$

---

[10] This approach is taken by Mirrlees (1976) and Holmström (1979). For each γ the distribution function $Y(M; \gamma)$ is given by $Y(M; \gamma) = H(c(M, \gamma))$, where H is the distribution function of c and $c(M, \gamma)$ is obtained by inverting $M = M(c, \gamma)$. The density function $y(M; \gamma)$ is then $y(M; \gamma) = \partial Y(M; \gamma)\, \partial M = h(c(M, \gamma))\, |\partial c/\partial M|$. If the density function $h(c)$ is defined on the interval $(\underline{c}, \bar{c})$, the density y is defined on $(M(\underline{c}, \gamma), M(\bar{c}, \gamma))$ and will be assumed to equal zero at its bounds. Additionally, y is assumed to be twice differentiable in γ.

where argmax denotes the set of arguments that maximize the succeeding function.

Specification of the conditions characterizing the optimal γ satisfying (17) is difficult because of the complexity of the functions ρ and y. The first-order condition is

$$\int \rho(M)\, y_\gamma(M)\, dM = 0, \tag{18}$$

since only the distribution of M depends on γ. The fuel-capital ratio satisfying (18) will be a local optimum if

$$\int \rho(M)\, y_{\gamma\gamma}(M)\, dM < 0,$$

but in general it is difficult to determine if for a given $\rho(M)$ the γ satisfying (18) is a global maximum. This issue as well as the existence of a solution to the regulator's program are addressed in Appendix C.

If (18) does characteristic the firm's choice of technology, replacing (17) with (18) and optimizing pointwise yields the first-order condition for $\rho(M)$

$$S_\rho^* + \hat{\lambda} + \hat{\psi} y_\gamma / y = 0, \tag{19}$$

where $\hat{\lambda}$ and $\hat{\psi}$ are multipliers corresponding to (16) and (18), respectively. The optimal price functions $\hat{p}(M)$ and $\hat{T}(M)$ satisfy the following conditions analogous to (7) and (8)

$$-S_p/\pi_p = \hat{\lambda} + \hat{\psi} y_\gamma / y, \tag{20}$$

$$-S_T/\pi_T = \hat{\lambda} + \hat{\psi} y_\gamma / y \qquad \text{if} \quad \hat{T} > 0. \tag{21}$$

Comparing these conditions with (7) and (8) indicates that in contrast to the full-information solution the functions $\hat{p}(M)$ and $\hat{T}(M)$ depend on the function y and hence on $h(c)$. Also, the first-best price formulas will not be utilized unless $\hat{\psi} y_\gamma / y$ is zero, so in general only a second-best solution is obtained.

Recalling that the regulator could force the full-information solution γ^* if γ were observable, the term y_γ / y in (20) and (21) is seen to be the derivative of the log-likelihood function where γ is viewed as an unknown parameter.[11] The magnitude of y_γ / y indicates how strongly the regulator is able to infer from the observation of M that the firm deviated from the first-best technology. The greater is $|y_\gamma|/y$ the more likely it is that the technology chosen by the firm differs from the first-best technology and hence the greater is the required deviation from the full-information price functions.

---

[11] This interpretation is developed by Holmstrom (1979).

The multiplier $\hat{\psi}$ may be interpreted as the "cost" of deviating from the full-information price functions in order to induce the firm to choose a technology that is more in the interests of consumers than is the technology that would be chosen by the firm if the regulator used price functions $p(M)$ and $T(M)$ satisfying (7) and (8). The multiplier $\hat{\psi}$ is characterized in the following proposition.

Proposition 4. *If the second-order condition for* (18) *is satisfied, the sign of* $\hat{\psi}$ *is the same as the sign of*

$$\int S^* y_\gamma \, dM = \int S^* \left[h' \cdot \frac{\partial c}{\partial \gamma} \frac{\partial c}{\partial M} + h \frac{\partial^2 c}{\partial M \, \partial \gamma} \right] dM.$$

Proof. The adjoint equation for ψ is obtained by maximizing the program in (15), (16), and (18) with respect to γ which yields

$$\int S^* y_\gamma \, dM + \hat{\psi} \int p(M) \, y_{\gamma\gamma} \, dM = 0.$$

The last term is $\hat{\psi}$ multiplied by the second-order condition for γ, which is negative since $\hat{\gamma}$ is optimal. ▮

The term $\int S^* y_\gamma \, dM$ represents the regulator's ex ante preference for a variation in γ with a negative (positive) sign indicating that the regulator prefers a fuel-capital ratio that is lower (higher) than that chosen by the firm. The multiplier $\hat{\psi}$ is thus negative (positive).

To further investigate the optimality of the price functions $\hat{p}(M)$ and $\hat{T}(M)$ characterized by (20) and (21), note that the sign of $S_p^* = S_p/\pi_p$ is the opposite of the sign of $(\hat{\lambda} + \hat{\psi} y_\gamma / y)$. Since $S_p = -Q$, $(\hat{\lambda} + \hat{\psi} y_\gamma / y) < 0$ implies that the price p satisfying (20) is such that $\pi_p < 0$ and similarly $\pi_T < 0$ if $T > 0$ satisfies (21). In this case a decrease in p (or T) would increase both profit and consumer surplus until the point at which $\pi_p = 0$.[12] The second-best price adjustment formulas thus satisfy (20) and (21) if $\hat{\lambda} + \hat{\psi} y_\gamma / y \geqslant 0$, and if $\hat{\lambda} + \hat{\psi} y_\gamma / y < 0$, the optimal prices are such that $\pi_p = 0$ and $\pi_T = 0$. In the second-best solution marginal profit is thus nonnegative which is the counterpart of Proposition 1.

When a two-part tariff is optimal, an expression analogous to (11) results

$$\hat{p} \left[1 + \left(1 - \frac{1}{\hat{\lambda} + \hat{\psi} y_\gamma / y} \right) \left(\frac{1 - N(z^0) \, s(z^0)}{\varepsilon} \right) \right] = \varphi' M. \tag{22}$$

---

[12] For a one-part price the second-order condition corresponding to p is $-Q_1 +$ $+ \hat{\psi} y_\gamma / y) \, \pi_{pp} < 0$. Concavity of π in p and $\hat{\lambda} + \hat{\psi} y_\gamma / y < 0$ implies that the p satisfying (23) is ocal minimum rather than a maximum.

The characterization of the second-best pricing policy is less conclusive than that in Proposition 2 for the first-best policy, since the term $(\hat{\lambda} + \hat{\psi} y_\gamma/y)$ depends on M and cannot be related to a returns to scale measure. The following characterization which obtains directly from (21) and (22) can be given, however.

Proposition 5. *If* $\hat{\lambda} + \hat{\psi} y_\gamma/y > (=)(<)\, 1$, $\hat{T}(M) > (=)(=)\, 0$ *and* $\hat{p}(M) > (=)(<)\, \varphi' M$.

The second-best solution may thus involve the use of a two-part price structure for some M and a one-part structure for other values of M. Such a pricing policy is necessary to obtain the optimal balance between risk sharing and incentive considerations.

For the case of a one-part price structure the expression in (22) may be rearranged to obtain

$$\frac{\hat{p} - \varphi' M}{\hat{p}} = \left(\frac{1}{\hat{\lambda} + \hat{\psi} y_\gamma / y} - 1 \right) \frac{1}{\varepsilon}. \tag{23}$$

The left-side is the proportional deviation from marginal cost pricing and depends on the elasticity of demand and the magnitude of $(\hat{\lambda} + \hat{\psi} y_\gamma/y)$. The expression in (23) is analogous to that for the second-best deviation from marginal cost pricing (see Baumol and Bradford) and indicates that

$$\hat{p} > (=)(<)\, \varphi' M \Leftrightarrow \hat{\lambda} + \hat{\psi} y_\gamma/y < (=)(>)\, 1.$$

The proportional deviation from marginal cost pricing is thus greater the greater is the derivative y_γ/y of the log-likelihood function and hence the regulator must make a greater compromise of the first-best sharing of the risk due to the factor price uncertainty as $|y_\gamma|/y$ increases.

To determine the relationship between the unit price and the unit cost M, consider the case in which a one-part price structure is utilized. For $\hat{\lambda} + \hat{\psi} y_\gamma/y > 0$ implicit differentiation of (19) yields an expression analogous to (12)

$$\hat{p}'(M) = -\left[(\hat{\lambda} + \hat{\psi} y_\gamma/y)(-\varphi' Q_1) + \pi_p \hat{\psi} \frac{\partial (y_\gamma/y)}{\partial M} \right] \Big/ D,$$

where D is the second-order condition which is negative when $\hat{\lambda} + \hat{\psi} y_\gamma/y > 0$. Consequently, a sufficient but not a necessary condition for $\hat{p}(M)$ to be an increasing function of M is that (y_γ/y) be decreasing (increasing) in M when $\hat{\psi} < (>)\, 0$, since regulation is effective $(\pi_p > 0)$ when $\hat{\lambda} + \hat{\psi} y_\gamma/y > 0$.[13]

---

[13] The second-best profit $\hat{\beta}(M)$ may be written analogously to (13) as $\hat{\beta}(M) = \varphi(1/\eta - 1)M + (1/(\hat{\lambda} + \hat{\psi} y_\gamma/y) - 1)\, Q^2(1 - N(z^0)\, s(z^0))/Q_1 + \hat{T} N(z^0)$. Even with constant returns to scale $(\eta = 1)$ the second-best profit function can be increasing, decreasing, or constant for various ranges of M.

When only the cost M is observable, a second-best solution results, and in this case there is an opportunity to obtain Pareto superior adjustment formulas by basing those formulas on other information that the regulator may obtain. For example, an electric utility must make decisions regarding alternative generation technologies, technologies for transmission and distribution, maintenance and servicing policies, employment of labor of various skills, etc., and the regulator is unlikely to be able to observe all of these decisions. When a regulator cannot monitor all of the components of the technology decision, there may be data collection activities that can provide useful information. For example, accounting data can provide information regarding fuel inputs and their costs, engineering studies can be conducted regarding line losses, heat rates, etc., and econometric studies can provide data on input decisions and their efficiency. Similarly, information could be obtained regarding factor prices of fuel, the sulfur content of alterative fuel supplies, transportation costs, etc.

Holmström (1979) has shown that if information represented by the ealization ω of a random variable is such that $y_\gamma(M, \omega; \gamma)/y(M, \omega; \gamma)$ epends on ω, where $y(M, \omega; \gamma)$ is the joint density of M and ω, then the ptimal second-best price functions depend on ω as well as on M. Adjustment mechanisms would thus be expected to depend on information egarding the firm and its economic environment in addition to incurred cost ata. This information is, however, rarely costless, so the gains from the nformation must be compared with the associated costs. While the optimal nformation acquisition decision will not be investigated here, this analysis oes provide a justification for the data collection activities of a regulatory commission. These issues will be considered further in the final section.

5. THE CHOICE OF A FUEL SUPPLY

An electric utility has the opportunity to purchase fuels from a variety of ources and may expand the range of choice by considering the conversion f its fossil-fuel-fired boilers from oil and gas to coal, for example. In a study f the economics of interfuel substitution Atkinson and Halvorsen (1976) ed a translog ex post cost function with capital and labor as fixed factors test for the degree of substitution among coal, oil, and gas as inputs, and oncluded "that steam electric power generation is characterized by bstantial ex post interfuel substitution..." Even if the choice among altertive fuels is restricted by their availability, utilities can choose among lternative sources of coal supply, for example. In an empirical study of fuel djustment mechanisms Gollop and Karlson (1978) concluded that in 1971 nly in the "coal belt" states was there a positive relationship between neration costs and FAC's. The authors speculate that this may be due to the presence of holding companies that own both coal mines and electric utilities or because of "less tight regulation." The objective of this section is to consider the design of a fuel adjustment mechanism to deal with the potential incentive problem associated with the fuel supply decision.

The model to be investigated includes both the case of perfect substitutes and the case in which conversion can be made from one fuel source to another. Perfect substitutes would normally have the same market price, but market imperfections such as those created by government regulations can cause the prices to differ. While the firm can be forced to choose the efficient fuel supply if the regulator can observe all available factor prices, the regulator is unlikely to have that degree of observability. Instead, he is more likely to rely on data regarding the firm's incurred costs obtained through accounting reports. In order to provide a comparison with the first-best policy, the regulator will be assumed to observe the incurred unit fuel cost and not the factor prices of all potential supplies. In order to focus on the fuel mix decision, the fuel-capital ratio will be assumed to be fixed, and only a one-part price structure will be considered.

To represent the fuel mix decision in a simple manner, the firm will be assumed to have a choice between two fuel sources, the first with uncertain factor price c and the second with a certain factor price c^+. The fuel decision will be assumed to be made on an ex post basis after the realization of c which corresponds to the use of short-term supply contracts or to spot purchases. The first fuel is taken to be the current supply source, so switching to the second fuel may involve conversion costs. The resulting unit cost of fuel will be denoted by $C(c, c^+, \delta)$, where $\delta \in [0, 1]$ denotes the fuel mix. Purchase of only the first fuel will be represented by $\delta = 1$, so $C(c, c^+, 1) = c$. Partial conversion to the second fuel is denoted by $\delta \in (0, 1)$ with $\delta = 0$ indicating complete conversion. The unit cost function will be assumed to be strictly increasing in c for $\delta \in (0, 1]$ and to be strictly convex in δ representing costs of conversion that increase at an increasing rate.

Since the fuel supply decision is made ex post, the efficient fuel supply minimizes the unit fuel cost C. Because of the observability problem, however, a profit-maximizing firm may have an incentive to purchase fuel inefficiently if its ex post profit $\rho(C)$ will be enhanced. If $\rho(C)$ is nonincreasing in C, the firm will not have an incentive to purchase an inefficient fuel mix, but otherwise an incentive for inefficiency is present. Consequently, the full-information profit function $\rho^*(C)$ given in (15) can be utilized when it is nonincreasing which includes the cases of constant returns to scale and constant price elasticity with a homogeneous production function.

When full-information profit function is increasing on some interval, the design problem faced by the regulator is to choose a function $\rho(C)$ that will provide the optimal balance between risk sharing and efficiency. To begin the characterization of the second-best profit function when an incentive problem is present, let $C^0 \equiv C(c, c^+, 0)$ be the cost with complete conversion to the second fuel. If the efficient fuel supply is chosen, the resulting unit fuel cost C is bounded above by C^0, so the regulator can specify $\rho(C) = -R$ for $C > C^0$, where R is a large penalty. This prevents the firm from choosing the inefficient fuel supply when the factor price c is high. This situation could correspond to the forced conversion from oil to coal when the factor price of oil rises to a particular level, for example. While such a penalty eliminates he incentive for inefficiency when c is high, the incentive problem remains or lower factor prices. For example, if $\rho(C)$ is a strictly increasing function ver $[0, C^0]$, the firm has an incentive to choose complete conversion to the econd fuel for any c and hence receive a profit $\rho(C^0)$ that is greater for all ctor price levels than the profit with the efficient fuel mix. In this case the utput price $p(C^0)$ must equal average cost $[\varphi(Q(p(C^0)))\, M(C^0, \gamma)/Q(p(C^0))]$ the profit constraint is to be satisfied. The regulator can thus use an verage cost pricing rule ($\rho(C) = 0$ for $C \leqslant C^0$), and the firm will be indifrent between the efficient and any inefficient fuel supply.

A more complex situation is illustrated by the profit function $\rho(c)$ shown Fig. 1. In this case the firm has an incentive to achieve the cost C_1 in der to earn the maximum possible profit. The set $\mathscr{C}_1$ of factor price levels ch that C_1 is attainable is

$$\mathscr{C}_1 = \{c \mid C(c, c^+, \delta) = C_1 \text{ for some } \delta \in [0, 1]\},$$

nce for any c the set of realizable unit costs is $[c, C^0]$. The firm has no

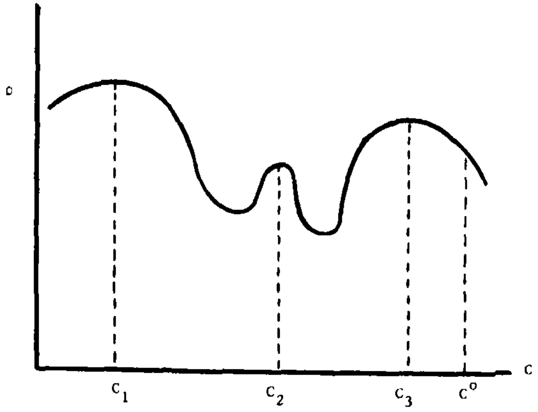


FIG. 1. The full-information profit function.

incentive to achieve the unit cost C_2 because C_3 is realizable, so C_3 will result on the set

$$\mathscr{C}_3 = \{c \mid C(c, c^+, \delta) = C_3 \text{ for some } \delta \in [0, 1]\} - \mathscr{C}_1.$$

For $c \notin \mathscr{C}_1 \cup \mathscr{C}_3$ the firm has the incentive to purchase the efficient fuel.

The regulator thus must select a profit function recognizing that only certain unit costs will be realized if $\rho(c)$ has increasing segments. The solution to this problem is difficult at best, and a complete characterization will not be given here. Instead, the optimal profit function in the class of nonincreasing functions that are differentiable on $[0, C^0)$ will be characterized. A member of this class is such that no incentive is created for inefficiency, and since $\rho(C) = 0$ for all C is a member of this class, the solution will be at least as good as using an average cost pricing rule. Letting $\hat{\delta}(c)$ denote the efficient choice of δ given a factor price c, the unit cost is $C = C(c, c^+, \hat{\delta}(c))$. The distribution of c thus induces a distribution of C, and let $v(C)$ denote the density function for $C < C^0$ with $v(C^0)$ denoting the probability that $C \geqslant C^0$.[14] The design problem can then be stated as

$$\begin{aligned} &\max_{\rho(C)} \int S^*(\rho(C), C)\, v(C)\, dc \\ &\text{subject to} \int \rho(C)\, v(C)\, dC \geqslant 0 \\ &\rho'(C) \leqslant 0 \end{aligned} \tag{24}$$

The following proposition characterizes the optimal solution $\bar{\rho}(C)$ to this program for the class of continuous, piece-wise differentiable functions.

---

[14] With the efficient fuel supply, the unit fuel cost $C = c$ for $C \in [0, c^+]$ and is $C = (c, c^+, \hat{\delta})$ for $C \in (c^+, C^0]$ where $\hat{\delta}(c)$ satisfies $C_\delta = 0$. The distribution function $V(C)$ is then

$$\begin{aligned} &= H(C) && \text{if } C \leqslant c^+ \\ V(C) &= H(c(C, c^+, \hat{\delta})) && \text{if } c^+ < C < C^0 \\ &= 1 && \text{if } C \geqslant C^0, \end{aligned}$$

where $c(C, c^+, \hat{\delta})$ is obtained by inverting $C = C(c, c^+, \hat{\delta})$ given that $\hat{\delta}(c)$ is strictly increasing on (c^+, C^0). The density function $v(C)$ corresponding to $V(C)$, for $C < C^0$, is

$$\begin{aligned} v(C) &= h(C) && \text{if } C \leqslant c^+ \\ &= h(c(C, c^+, \hat{\delta})) \left| \frac{\partial c}{\partial C} \right| && \text{if } c^+ < C < C^0, \end{aligned}$$

and at $C = C^0$ there is a mass point with mass $v(C^0) = \Pr(c \geqslant C^0)$.

PROPOSITION 6. *Let $\bar{\rho}(C)$ denote the optimal solution to the program in* (24). *If $\bar{\rho}'(C) < 0$ on an open interval, the optimal profit function $\bar{\rho}(C)$ and price function $\bar{p}(C)$ on that interval satisfy* (6) *and* (7), *respectively, with λ^* replaced by the multiplier $\bar{\lambda}$ from the program in* (24). *If $\bar{\rho}'(C) = 0$ on an pen interval, the optimal profit is a constant $\bar{\rho}$ (which depends on the nterval) and $\bar{p}(C)$ satisfies*[15]

$$-(1 + S_{\rho}^{*}(\bar{\rho}, C))\, Q(\bar{p}(C)) - S_{\rho}^{*}(\bar{\rho}, C)(\bar{p}(C) - \varphi' M)\, Q_1(\bar{p}(C)) = 0.$$

Proof. Viewing (24) as a control problem with $\rho(C)$ the state variable nd $\rho'(C)$ the control, the necessary conditions may be written as

$$\begin{aligned} \bar{\xi}'(C) &= -(S_{\rho}^{*} + \bar{\lambda})\, v(C), \\ \bar{\xi}(C)\, \bar{\rho}'(C) &= 0, \end{aligned} \tag{25}$$

here $\bar{\xi}(C)$ is the multiplier associated with the constraint $\rho'(C) \leqslant 0$. Conseuently, $\bar{\rho}'(C) < 0$ on an open interval implies $\bar{\xi}(C) = 0$ and $\bar{\xi}'(C) = 0$, so 5) reduces to (6), which implies (7), on that interval. If $\bar{\xi}(C) > 0$ on an pen interval, $\bar{\rho}'(C) = 0$ and $\bar{\rho}(C)$ is a constant $\bar{\rho}$. Then, evaluating $_{p} = S_{\rho}^{*}\pi_{p}$ implies the desired result. ∎

he profit function $\bar{\rho}(C)$ characterized in Proposition 6 is illustrated in g. 2. The losses compared to the full-information solution result both cause $\bar{\rho}(C)$ is not permitted to increase and because $\bar{\rho}(C) > \rho^*(C)$, at least n some interval, in order to compensate the firm for reduced profits on tervals such that $\rho^{*\prime}(C) > 0$.

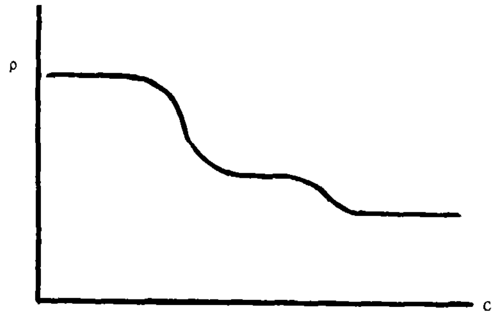


FIG. 2. The second-best profit function.

---

[15] It may be possible to improve on this solution by permitting discontinuities or by owing inefficiency over some range of factor prices. Unfortunately, a general solution to the ign problem is not yet available for the case in which the firm can make ex post decisions.

6. DISCUSSION

While there is no conclusive evidence that regulatory price adjustment mechanisms are preferable to administrative rate review procedures, their widespread use in the electric utility industry indicates that such mechanisms have certain perceived advantages in terms of matching revenues and costs, in lessening the deterioration of earnings, and in risk sharing. If such mechanisms are to be employed, the full-information formulas involve gains to risk sharing resulting from deviating from average cost pricing (except, of course, for the case of constant returns to scale). When observability is incomplete, however, the predictability of the price changes under regulatory adjustment mechanisms can pose incentive problems that result in technical inefficiency and increased cost. Firms may have an incentive to favor factor inputs that are anticipated to lead to output price increases and may choose factor input supply sources that are cost enhancing. These incentive problems can be present because of the inability of a regulatory commission to observe the decisions made by the firm and certain exogenous variables such as factor prices. Reliance on accounting reports for monitoring the actions of a firm can contribute to the incentive problems, since only incurred cost data are reported. Such reliance requires a regulatory commission to deviate from the full-information formulas in order to mitigate potential incentives for inefficiency. This implies that the regulator may be justified in obtaining a capability, perhaps limited, to monitor exogenous variables and to develop the technical expertise to evaluate the efficacy of investment and supply decisions. Information obtained in this manner will be useful in designing adjustment mechanisms if that information is informative in Holmström's sense.

In the context of the model considered here, the observation of factor prices and the ability to ascertain conversion costs would be sufficient to eliminate potential fuel supply inefficiencies. When only incurred costs can be observed, however, it is necessary to design price adjustment formulas so that the ex post profit of the firm is a decreasing function of unit costs or to tolerate the chance of endogenously increased fuel costs. Instead of deviating from the first-best risk sharing function in order to induce the efficient supply source, legislation that would require conversion to lower cost fuels or would impose penalties to induce conversion may be a better solution when the relative price of fuel is anticipated to increase.

The potential incentive to bias the choice of technology when a regulatory price adjustment mechanism is employed provides a justification for a commission's function of certification of investments. This, of course, requires a significant expenditure to develop the technical expertise necessary to be able to evaluate investment alternatives, but some staff capabilities in this area may be warranted. The commission could additionally conduct cross-sectional studies of electric utilities to compare performance of a particular firm with that of others facing similar conditions. When such monitoring provides informative data regarding the firm's decisions or factor prices, Pareto superior adjustment mechanisms can in principle be developed.

APPENDIX A: PROOF OF PROPOSITION 2

First, it will be shown that $\lambda^* > 1$ when returns to scale are increasing. If $\lambda^* \leqslant 1$, then evaluated at $T = 0$, (8) implies that $S_T + \lambda^* \pi_T \leqslant 0$ and the concavity of $(S + \lambda\pi)$ implies that $T(c) = 0$ is optimal for all c. Then, Q^+ and π_{t_0} are zero, and (7) implies that $p = \varphi' M \leqslant 0$ for all c. Pricing less than or equal to marginal cost with increasing returns to scale implies that profit is negative for all c, so $E\pi < 0$. This contradiction implies that $\lambda^* > 1$.

To show that $T^*(c) > 0$ and $p^*(c) > \varphi' M$ for all c, note that for $\lambda^* > 1$, $S_T + \lambda^* \pi_T > 0$ in (8) when evaluated at $T = 0$, so $T^*(c) > 0$ is optimal for all c. Since $T^*(c) > 0$, (10) holds. When $T^* = 0$, $Q^+ = 0$, and (7) reduces to (10), so (10) holds for all cases. Since $Q - N(z^0)\, Q^+ > 0$, $\lambda^* > 1$ implies that $p^*(c) > \varphi' M$ for all c.

For the case of constant returns to scale, the proof proceeds by first showing that $\lambda^* = 1$. If $\lambda^* > 1$, then $T(c) > 0$ from above. Since $(Q - N(z^0)\, Q^+) > 0$, the expression in (10) and $\lambda^* > 1$ imply that $p > M$ with constant returns to scale ($\varphi' = 1$). Consequently, $\pi(c) > 0$ for all c, so $E\pi(c) > 0$ which implies $\lambda^* = 0$. This contradicts $\lambda^* > 0$, so $\lambda^* \not> 1$. If $\lambda^* < 1$, then $T = 0$ from above, and (10) implies that $p < \varphi' M = M$. Consequently, $E\pi(c) < 0$, which contradicts the assumption that $\lambda^* < 1$, so $\lambda^* = 1$. This implies that $T^*(c) = 0$ for all c and from (10) that $p^*(c) = M$ for all c.

To show that $\lambda^* < 1$ with decreasing returns to scale, assume the contrary. Then (10) implies that $p^* \geqslant \varphi' M$, so with price at least as high as marginal cost for all c, $E\pi > 0$. Consequently, $\lambda^* < 1$. Then, $T^*(c) = 0$ and from (10) $p^*(c) < \varphi' M$ for all c. ∎

APPENDIX B: CHARACTERIZATION OF THE FULL-INFORMATION PRICE FORMULAS

To further characterize the full-information price formulas, rewrite (10) as:

$$p^* = \varphi' M + p^* \left(\frac{1/\lambda^* - 1}{\varepsilon}\right) (1 - N(z^0)\, s(z^0)) \tag{B1}$$

$$= \frac{\varphi}{Q} M + \varphi M \left(\frac{1}{\eta} - 1\right) \Big/ Q + p^* \left(\frac{1/\lambda^* - 1}{\varepsilon}\right) (1 - N(z^0)\, s(z^0)), \tag{B2}$$

where $\eta \equiv \varphi/(\varphi' Q)$ is the returns to scale measure ($\eta >$ (=)(<) 1 indicating increasing (constant) (decreasing) returns). The expression in (B1) gives the deviation from marginal cost ($\varphi' M$) pricing, while the expression in (B2) gives the deviation from average cost pricing. If ex post regulation were utilized, average cost pricing would be required, since a fair-return constraint ($\pi = 0$) would be imposed for all c. With ex ante regulation the regulator is able to to take advantage of the opportunity to deviate from average cost pricing and hence to increase expected consumer surplus while still providing the firm with an ex ante fair return.

The deviation from average cost pricing consists of two parts represented by the second and third terms in (B2). The first of those terms represents a divergence from average cost pricing in response to the returns to scale of the firm, while the second term represents a compensating effect whose magnitude depends on the demand function. With increasing returns to scale the term $\varphi M(1/\eta - 1)/Q$ is negative reflecting an adjustment in the direction of marginal cost. The last term in (B1) and (B2) is positive, so the unit price could be above average cost for some for c when a two-part pricing structure is utilized. The following table summarizes the characterization of the first-best price functions.

				Deviations from average cost	
Returns to scale	λ^*	$T^*(c)$	$p^* - \varphi' M$	Returns to scale effect $\varphi M(1/\eta - 1)/Q$	Demand effect $(p^*/\varepsilon) \cdot (1/\lambda^* - 1) \cdot (1 - Ns)$
$\eta > 1$	>1	+	+	−	+
$\eta = 1$	$=1$	0	0	0	0
$\eta < 1$	<1	0	−	+	−

APPENDIX C: COMMENTS ON THE SECOND-BEST SOLUTION

The existence of a solution to the program in (15)–(17) is an open question for the general class of functions $\rho(C)$. Holmström (1977, pp. 215–217) has provided an existence theorem for this class of problem and shown that if attention is restricted to the class of bounded, equi-continuous functions, a solution exists. Existence will be assumed here with bounds imposed on the price and profit functions as needed for existence.

If a solution exists to the regulator's program in (15)–(17), the use of (18) in place of (17) can cause difficulties in the sense that a function $\rho^+(C)$

satisfying (19) may correspond to a γ^+ satisfying (18) that is only a local maximum for the firm. If the regulator then implemented $p^+(C)$ on the anticipation that the firm would choose γ^+, the firm upon globally maximizing could choose a γ different from γ^+ resulting in a welfare level hat is quite different from the second-best optimum. The "solution" charac- erized here is assumed to be the second-best optimum.

Mirrlees (1974) has shown in the context of related problems that while a olution may not exist in the class of unbounded functions, the first-best olution can, however, be approximated arbitrarily closely by imposing penalties that would have a negligible effect on the expected profit of the firm while inducing the firm to choose a technology arbitrarily close to the first- est technology. Whether this is possible in general is as yet unclear. The use f penalties can be illustrated from a somewhat different perspective by noting that the support of the distribution of the unit cost M depends on the el-capital ratio. When this is the case, it is possible to detect certain deviations from the first-best fuel-capital ratio. If γ^* is the first-best chnology and the support of c is $[\mathbf{c}, \infty)$, for example, the support of M at is $[M(\mathbf{c}, \gamma^*), \infty)$. If p^0 denotes a (low) penalty price, the regulator could ilize a priced function $p(M)$ of the form

$$\begin{aligned} p(M) &= p^+(M) \qquad \text{if} \quad M \geqslant M(\mathbf{c}, \gamma^*) \\ &= p^0 \qquad\qquad \text{if} \quad M < M(\mathbf{c}, \gamma^*) \end{aligned}$$

nce $M(\mathbf{c}, \gamma^*)$ is presumably decreasing in γ for low factor prices, there is a positive probability that the price p^0 would be implemented if the firm were choose $\gamma \geqslant \gamma^*$. By choosing p^0 sufficiently low, it may be possible to compel the firm to choose a fuel-capital ratio no greater than γ^*. If there is incentive to choose too great a fuel-capital ratio because of an anticipated crease in the price of fuel, such a penalty will be effective in promoting chnical efficiency.

ACKNOWLEDGMENTS

The authors would like to thank Bengt Holmström for his helpful comments.

REFERENCES

S. E. ATKINSON AND R. HALVORSEN, Interfuel substitution in steam electric power generation, *J. Pol. Econ.* **84**(Oct. 1976), 959–978.

E. E. BAILEY, Innovation and regulation: A reply, *J. Public Econ.* 5 (1976), 393–394.

D. P. BARON AND R. R. DE BONDT, Fuel adjustment mechanisms and economic efficiency, *J. Industrial Econ.* 27 (March 1979), 243–261.

4. W. J. Baumol and D. R. Bradford, Optimal departures from marginal cost pricing, *Amer. Econ. Rev.* **60** (June 1970), 265–283.
5. P. A. Diamond and J. A. Mirrlees, Optimal taxation and public production II: Tax rules, *Amer. Econ. Rev.* **61** (1971), 261–278.
6. Edison Electric Institute, "October 3, 1974 Questionnaire on Fuel Adjustment Revenue," New York.
7. G. R. Faulhaber and J. C. Panzar, "Optimal Two-Part Tariffs with Self-Selection," Working Paper, Bell Laboratories, Holmdel, New Jersey, 1977.
8. F. M. Gollop and S. H. Karlson, The impact of the fuel adjustment mechanism on economic efficiency, *Rev. Econ. Statist.* **60** (November 1978), 574–584.
9. B. Holmström, "On Incentives and Control in Organizations," unpublished dissertation, Stanford University, 1977.
10. B. Holmström, Moral hazard and observability, *Bell J. Econ.* **10** (Spring 1979), 74–91.
11. H. Hotelling, The general welfare in relation to problems of taxation and of railway and utility rates, *Econometrica* **6** (July 1938), 224–269.
12. P. Joskow, Inflation and environmental concern: Structural change in the process of public utility regulation, *J. Law Econ.* **17** (October 1974), 291–327.
13. H. E. Leland and R. A. Meyer, Monopoly pricing structures with imperfect discrimination, *Bell J. Econ.* **7** (Autumn 1976), 449–462.
14. J. A. Mirrlees, The optimal structure of incentives and authority within an organization, *Bell J. Econ.* **7** (Spring 1976), 105–131.
15. J. A. Mirrlees, Notes on welfare economics, information and uncertainty, *in* "Essays on Economic Behavior Under Uncertainty" (M. S. Balch, D. L. McFadden, and S. Y. Wu, Eds.), pp. 243–258, North-Holland, Amsterdam, 1974.
16. Naruc, Automatic adjustment clauses revisited, Economic Paper No. IR, "86th NARUC Annual Convention," 1974, pp. 313–318; also reprinted in U. S. Senate (1977), pp. 88–98.
17. New Mexico Public Service Commission, "Public Service Company of New Mexico," Case No. 1196, April 22, 1975. (8 *PUR* 4th, pp. 113–139).
18. Y. Ng and M. Weisser, Optimal pricing with a budget constraint—the case of a two-part tariff, *Rev. Econ. Studies* **41** (July 1974), 337–345.
19. W. Y. Oi, A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly, *Quart. J. Econ.* **85** (February 1971), 77–96.
20. K. W. S. Roberts, Welfare considerations of nonlinear pricing, *Econ. J.* **89** (March 1979), 66–83.
21. R. Schmalensee, Option demand and consumer's surplus: Valuing price change under uncertainty, *Amer. Econ. Rev.* **62** (December 1972), 813–824.
22. M. Spence, Nonlinear prices and welfare, *J. Public Econ.* **8** (1977), 1–18.
23. U. S. House of Representatives, Subcommittee on Oversight and Investigations, "Electric Utility Automatic Fuel Adjustment Clauses," Government Printing Office, Washington, October, 1975.
24. U. S. Senate, "Electric and Gas Utility Rate and Fuel Adjustment Clause Increases, 1976," prepared by the Congressional Research Service for the Subcommittee on Intergovernmental Relations and the Subcommittee on Reports, Accounting and Management, Government Printing Office, Washington, July, 1977.
25 R. D. Willig, Consumer's surplus without apology, *Amer. Econ. Rev.* **66** (September 1976), 589–597.

URNAL OF ECONOMIC THEORY **24**, 95–111 (1981)

Equilibrium in Dynamic Models with an Infinity of Agents

CHARLES A. WILSON

University of Wisconsin, Madison, Wisconsin 53706

Received August 18, 1980

1. INTRODUCTION

In this paper, I investigate the existence of a competitive equilibrium in a ynamic model of pure exchange with a countable number of agents and a untable number of goods. The assumptions I employ are of the same level generality as those generally used to establish the existence of equilibrium classical competitive models. The commodity space is represented by the t of all non-negative sequences which contains the endowment of each ent. The preferences of each agent are assumed to satisfy a convexity quirement and to be continuous in the product topology. There is free dis- sal and the economy must satisfy a certain resource relatedness condition. der these assumptions, I then establish the existence of a competitive uilibrium in two special cases: (1) when the endowment of each agent is sitive for only a finite number of goods; (2) when the aggregate dowment of some finite subset of agents is greater than some fixed fraction the economy-wide endowment in each period. In all other cases, the most t I can establish is the existence of a special kind of equilibrium with nsfer payments. Specifically, I show that a set of market clearing prices l exist if we permit those agents endowed with a positive quantity of an nite number of goods to augment the value of their endowment by a finite -negative amount.

he existence of a competitive equilibrium when there is only a finite set agents has already been demonstrated by Bewley [3].[1] Recently Balasko Shell [1] and Balasko *et al.*, [2] have established the same result for e exchange economies composed of an infinite number of finite-lived nts. Although the assumptions which I adopt in this paper are not quite general in every respect as the assumptions used by either of these hors, it is possible to modify my assumptions so that the theorem onstrated here implies both Bewley's theorem (for discrete-good, pure-

---

[1] Bewley also includes production. Moreover, his results are demonstrated for any sigma- e measure space of commodities.

0022-0531/81/010095-17$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

exchange economies) and the theorems of Balasko *et al.* In addition, it covers the case where there is an infinite number of infinite-lived agents or perhaps a mixture of both infinite-lived and finite-lived agents.

The basic method of proof is to consider a sequence of subeconomies with a finite set of agents and a finite number of goods, establish the existence of a competitive equilibrium in each of these subeconomies, and then demonstrate that a limiting allocation and price is an equilibrium for the full economy. This is essentially the same idea as was used by Bewley. There are, however, some critical points of departure. These differences generally result from the fact that the equilibrium price systems for the class of models considered here cannot always be represented as elements of the dual to the commodity space. In particular, as is already known from the study of simple models of overlapping generations, the equilibrium price sequence need not assign a finite value to the aggregate endowment. In these cases, the usual implications of Walras' Law may no longer be valid, and, as a consequence a competitive equilibrium may not exist.

The assumptions which I use to ensure the existence of equilibrium in each of the finite agent subeconomies are based on the assumptions used by McKenzie [7]. His paper contains a synthesis of the contributions of several authors over the last decade and provides a very general existence result for finite good, finite agent economies. For my purposes, it has been necessary to strengthen some of these assumptions in order to ensure the necessary properties for the infinite agent economy. In some cases, however, restrictions have been introduced merely to simplify the argument and the statement of the results. Although I have generally tried to indicate how these restrictions might be relaxed, it seemed preferable to present an explicit proof of a more general theorem in a model that also includes production.

The paper is organized as follows. In Section 2, I outline the basic model and list the assumptions to be used. Section 3 contains a brief discussion of prices and deals with the problem of defining the budget set for prices which assign an infinite value to an agent's endowment. Section 4 contains the heart of the paper. I establish here the existence of a competitive equilibrium with positive transfer payments and also establish some conditions under which a competitive equilibrium will exist. In Section 5, I present the analogue of the first welfare theorem. Using the standard argument of Debreu [5], I show that if the value of the aggregate endowment is finite at the equilibrium price, then the competitive allocation is Pareto optimal. I then show that this condition will be satisfied when a significant fraction of the aggregate endowment is owned by only a finite number of agents. In the next section, I re-examine the role of some of the assumptions. The need for a resource relatedness assumption is emphasized and problems which arise in relaxing free disposal are discussed. I also indicate here how my results could be incorporated into the framework used by Bewley.

The paper concludes with an example of an economy for which there is no competitive equilibrium but for which there is an equilibrium if some agent is given a positive transfer payment. I argue that this can be interpreted as an example where a sequence of market clearing prices will exist only if some agents are endowed with a positive amount of Samuelson's money [8].

2. THE MODEL

We begin with the standard formulation of a pure exchange economy. Let A be a set of agents, and, for each agent α, let X_α be his consumption set, $(>)_\alpha$ his preference relation over X_α, and w^α his endowment. Then we may define an *economy* $\mathscr{E}$ as a set $(X_\alpha, (>)_\alpha, w^\alpha)_{\alpha \in A}$. An *allocation* for economy is then a set $(x^\alpha)_{\alpha \in A}$ with $x^\alpha \in X_\alpha$ for each $\alpha \in A$.

The set of goods will be indexed by the natural numbers N. In the context of a dynamic economy, we may think of there being an infinite number of discrete periods with a finite number of goods in each period. Therefore, we will assume:

ASSUMPTION I. *For each* $\alpha \in A$:

(i) $X_\alpha = R_+^\infty$,

(ii) $w^\alpha \in R_+^\infty$.

Condition (i) identifies the consumption set of each agent with the set of ll non-negative sequences; condition (ii) requires that the endowment of ch agent lie in this set. To the extent that we wish to incorporate finite-ved agents into the model, it might appear inappropriate to define each gent's consumption set over all non-negative sequences (particularly when e include the necessary continuity assumptions on preferences). Since we ill eventually introduce a free disposal assumption, however, this onvention implies no real restrictions. Also, it is not essential that w^α be on-negative. This restriction was introduced only to simplify the statement some of the assumptions and proofs.[2]

We turn next to the restrictions to be imposed on the preferences of dividual agents. For each $x \in X_\alpha$, let $P_\alpha(x) = \{z \in X_\alpha: z(>)_\alpha x\}$ and $^{1}(x) = \{z \in X_\alpha: x(>)_\alpha z\}$. $P_\alpha(x)$ represents the set of commodity bundles ich agent α strictly prefers to bundle x. Similarly, $P_\alpha^{-1}(x)$ represents those ndles to which x is strictly preferred by agent α.

---

[2] In fact it is sufficient to assume merely that X_i is convex and bounded below. However, statement of the irreducibility condition (Assumption VI following) becomes considerably re complicated.

ASSUMPTION II. *For each* $a \in A$ *and* $x \in X_a$:

(i) $P_a(x)$ *and* $P_a^{-1}(x)$ *are both open relative to* X_a *with respect to the product topology in* R^∞,[3]

(ii) *If* $z \in P_a(x)$, *then* $\lambda z + (1-\lambda)\, x \in P_a(x)$ *for* $0 < \lambda \leqslant 1$.

This assumption corresponds to Assumption (2″) in McKenzie [7], except for a slightly stronger continuity requirement, and a much more restrictive convexity condition.[4] The strong convexity requirement is needed to guarantee a specific form of local non-satiation for infinitely-lived agents. Note, however, that preferences need not be transitive or even asymmetric. The implications of continuity of preferences with respect to the product topology merit a few comments. First, for those agents whose preference orderings depend on only a finite number of goods, it amounts to no more than the usual continuity requirement in finite-dimensional commodity spaces. For agents who consume an infinite number of goods, it implies some degree of impatience. However, the assumption is satisfied by any utility function of the form $U(x) = \sum_{i=1}^{\infty} u(x_i, i)$, where $U(x)$ is well defined and finite for all x and $u(\cdot\, ; i)$ is continuous for all i. Moreover, since the set of goods is countable, Assumption II(i) is essentially no more restrictive than continuity with respect to the Mackey topology when the commodity space is restricted to l_∞^+, as in Bewley [3].[5] I will return to this point in Section 6.3.

We will also assume:

ASSUMPTION III. *For all* $a \in A$ *and* $x \in X_a$, *if* $z \in P_a(x)$ *and* $y \geqslant z$,[6] *then* $y \in P_a(x)$.

I introduce this restriction as a distinct assumption to emphasize that it is as much a statement about the technology as about individual preferences. It is no more restrictive than assuming free disposal with the corresponding implication that prices must be non-negative in equilibrium. Moreover, it is included only to simplify the statement of some of the theorems and proofs. I discuss briefly in Section 6.2 how the argument must be modified if this assumption is relaxed.

We turn next to the restrictions to be imposed on the economy as a whole. The first assumption specifies how large the set of agents can be.

---

[3] $x^t \to x$ in the product topology if $x_i^t \to x_i$ for all $i \in N$.

[4] McKenzie requires only that P_a be relatively open valued and lower semi-continuous and that $x \notin$ convex hull $P_a(x)$. In fact, this assumption would be sufficient here for agents who consume only a finite number of goods, if the irreducibility condition (Assumption VI) were suitably modified.

[5] If a sequence is uniformly bounded, it converges in the Mackey topology of l_∞ if and only if it converges in the product topology of R^∞.

[6] In what follows $x \geqslant 0$ means $x_i \geqslant 0$ for all i, $x > 0$ means $x \geqslant 0$ and $x \neq 0$, $x \gg 0$ means $x_i > 0$ for all i.

ASSUMPTION IV. *A is a countable set.*

This assumption introduces enough agents into the model so that models of overlapping generations can be included into the analysis. It also allows for possibly an infinite number of infinitely lived agents.

I will also suppose that the total endowment of each good is finite and define $w \equiv \sum_{\alpha \in A} w^\alpha$ to be the *aggregate endowment* of the economy.

ASSUMPTION V. $w \gg 0$.

This assumption merely requires that the aggregate endowment of each good be positive. It corresponds to Assumption (5) in McKenzie [7]. Given the specification of the consumption sets it implies that the aggregate endowment lie in the interioı of the aggregate consumption set.

In addition to Assumption V, we also need a condition which guarantees that the preferences and endowments of agents are sufficiently interconnected to generate a competitive equilibrium in an appropriate sequence of subeconomies. In Bewley's work, this requirement is met by assuming that each agent has a strictly positive endowment bounded away from zero and that preferences are weakly monotonic. Here, such an assumption is inappropriate since we wish to include the possibility that some of the agents are only finitely long lived. Balasko and Shell essentially assume that the preferences of any agent are strictly monotonic in any good with which they are positively endowed, and then assume that there is sufficient overlap in he endowments of agents. The assumption I will employ is weaker than either of these assumptions. It is based upon the concept of irreducibility used by McKenzie [7].[7] Here I use a simplified version appropriate to a model of pure exchange.

An economy is *irreducible* if, for any partition of A into two non-empty ubsets B_1 and B_2 and any allocation (x^α), there is a $\beta \in B_1$ and a $\leqslant \sum_{\alpha \in B_2} w^\alpha$ such that $z + x^\beta (>)_\beta\, x^\beta$.[8] Given the free disposal assumption II and the fact that consumption sets are identified with R^∞_+, this is quivalent to requiring that the aggregate endowment of any set of agents an always be added to a feasible consumption vector of some agent in the *omplement* set to make that agent better off. Note, in particular that, for any gent, it implies the existence of a point preferred to any feasible onsumption vector and also that the endowment of each agent must be nonro. It is not sufficient to assume that the economy as a whole is reducible, however. We also must be able to select an appropriate sequence of finite agent subeconomies which are irreducible. For this, the following assumption is sufficient.

---

[7] This ıs also the approach used by Balasko, Cass, and Shell [2].

[8] Gıven Assumption II(ii), this could be weakened to read: there is a $(z^\alpha)_{\alpha \in B_1}$; $z^\alpha \in R^\infty$ for $\alpha \in B_1$, such that $\sum_{\alpha \in B_1} z^\alpha \leqslant \sum_{\alpha \in B_2} w^\alpha$ and $x^\alpha \notin P_\alpha(z^\alpha + x^\alpha)$ for all $\alpha \in B_1$ with $+ z^\alpha (>)_\alpha\, x^\alpha$ for some $\alpha \in B_1$.

ASSUMPTION VI. (i) *The full economy* $\mathscr{E}$ *is irreducible, and,*

(ii) *For any finite set of agents* $B \subset A$, *there is a finite set* $C \subset A$ *which contains* B *for which the corresponding subeconomy* $(X_\alpha, (\succcurlyeq)_\alpha, w^\alpha)_{\alpha \in C}$ *is irreducible.*

3. PRICES

Before formally defining the concept of a competitive equilibrium, it is necessary to discuss what we mean by prices in such an economy. At the very least, a price system should be a sequence of prices—one for each individual good. If no other restrictions are introduced, however, the value of some feasible consumption bundles may be infinite. Consequently, in order to ensure that each agent's demand function is well defined, it appears that one of the following two approaches must be adopted. Either the price system (and possibly the consumption sets) must be restricted in a way that we consider only those prices which generate a well defined value for each feasible net trade; or, the budget set must be defined in a way that does not require that the value of all feasible net trades be well defined. Bewley [3], in his work on the existence of equilibrium with a finite number of agents, adopted the first approach. Since I am working with an infinite number of agents, I have found it more convenient to adopt the second approach. Thus, the budget set will be defined for all prices in R^∞. One of the implications of Theorem 1, however, is that the equilibrium price sequence must always lie in the set of prices which assign a finite value to the endowment of each agent.

Formally, a *price system* is a sequence $p \in R^\infty$. For p, $x \in R^\infty$, *define* $p \cdot x = \liminf_{T \to \infty} \sum_{t=1}^{T} p_t x_t$. Then for each $p \in R^\infty$, we may define $B_\alpha(p) = \{x \in X_\alpha : p(w^\alpha - x) \geqslant 0\}$ to be the *budget set* for consumer α. This seems to be a natural definition of the budget set for an agent in a dynamic economy with perfect capital markets. Assuming that goods are ranked in the order of the periods in which they are available, the budget constraints simply requires that there be no subsequence of net indebtedness. $(\sum_{t=1}^{T} p_t(x_t - w_t^\alpha))$, which is bounded above zero. This essentially rules out Ponzi games.[9] Given the definition of the budget set, we may then define for each

---

[9] Alternatively, we could have defined $p \cdot x = \limsup_{T \to \infty} \sum_{t=1}^{T} p_t x_t$ without affecting our results. In this case we require only that some subsequence of net indebtedness not converge to a positive number. Note that in neither case can p generally be regarded as a linear function.

$p \in R^\infty$, the value of agent α's *demand correspondence* to be $D_\alpha(p) = \{x \in B_\alpha(p)\colon P_\alpha(x) \cap B_\alpha(p) = \varnothing\}$. This corresponds to the usual definition. If a bundle x is to be demanded by agent α at price p, his budget set must contain x but not any bundle which α prefers to x.

4. EQUILIBRIUM

In this section, we establish some conditions under which a competitive equilibrium will exist. We will say that an allocation (x^α) is *attainable* if $\sum_{\alpha \in A} x^\alpha = w$. Then a price p and an allocation (x^α) is a *competitive equilibrium* (C.E.) for economy $\mathscr{E}$ if (x^α) is attainable and $x^\alpha \in D_\alpha(p)$ for all $\alpha \in A$. As *I* illustrate in Section 7, Assumptions I to VI are not sufficient to guarantee the existence of competitive equilibrium. It is possible, however, to establish the following result.

THEOREM 1. *There is a price $p > 0$ and an attainable allocation (x^α) such that for each $\alpha \in A$:*

(i) $0 < pw^\alpha \leqslant px^\alpha < \infty$;

(ii) *$x \in P_\alpha(x^\alpha)$ implies $p(w^\alpha - x) < p(w^\alpha - x^\alpha)$;*

(iii) *$pw^\alpha = px^\alpha$ if $w_i^\alpha > 0$ for only a finite number of i.*

Conditions (i) and (ii) establish what, borrowing from Bewley, we might call a competitive equilibrium with non-negative transfer payments. The allocation is attainable and any bundle which is preferred by some agent to his assigned bundle must cost more. However, the value of an agent's consumption bundle need not equal the value of his endowment. In order for the agent to purchase his endowment, it may be necessary to augment his purchasing power with a positive transfer payment. Only for the case of finite-lived owners does the theorem require that the consumption bundle of the agent actually lie in his consumption set.

Before presenting the proof of Theorem 1, let me first establish a set of conditions under which it implies the existence of a competitive equilibrium without transfer payments.

THEOREM 2. *If either:*

(i) *for each $\alpha \in A$, $w_i > 0$ for only a finite number of i, or*

(ii) *there is a finite subset $B \subset A$ and an $\varepsilon > 0$ such that $\sum_{\alpha \in B} w_i^\alpha > \sum_{\alpha \in A} w_i^\alpha$ for all but a finite number of i,*

then a competitive equilibrium exists.

Proof. *Consider the price p and allocation (x^α) of Theorem 1. If condition (i) is satisfied, then conditions (ii) and (iii) of Theorem 1 imply* that $x^\alpha \in D_\alpha(p)$ for all $\alpha \in A$. Hence $(p, (x^\alpha))$ is a competitive equilibrium. If condition (ii) is satisfied, then condition (i) of Theorem 1 implies $\infty > \sum_{\alpha \in B} pw^\alpha = p \sum_{\alpha \in B} w^\alpha > (1/\varepsilon)\, pw + k$ for some $k < \infty$. Therefore, $pw < \infty$. Also $p > 0$ and $w^\alpha, x^\alpha \geqslant 0$ for all $\alpha \in A$ implies $\sum_{\alpha \in A} px^\alpha = p \sum_{\alpha \in A} x^\alpha$ and $\sum_{\alpha \in A} pw^\alpha = p \sum_{\alpha \in A} w^\alpha = pw$. Then since $\sum_{\alpha \in A} x^\alpha = w$, we have $p \sum_{\alpha \in A} x^\alpha \leqslant pw < \infty$. Condition (i) of Theorem 1 then implies $p(w^\alpha - x^\alpha) = 0$ for all $\alpha \in A$. Q.E.D.

Theorem 1 will be established by considering a sequence of subeconomies consisting of only a finite number of agents whose consumption sets are restricted to be finite-dimensional subspaces of R^∞. Let $\{A_t\}_{t=1}^{\infty}$ be a sequence of finite subsets of A for which $A_t \subset A_{t+1}$ and $\bigcup_{t=1}^{\infty} A_t = A$. By Assumptions V and VI, we may choose A_t so that for each t, $\sum_{\alpha \in A_t} w_i^\alpha > 0$ for $i \leqslant t$ and the corresponding subeconomy $(X_\alpha, (>)_\alpha, w^\alpha)_{\alpha \in A_t}$ is irreducible. For each $\alpha \in A_t$, define $\tilde{w}_t^\alpha \in R_+^\infty$ by $(\tilde{w}_t^\alpha)_i = 0$ if $i \leqslant t$; $(\tilde{w}_t^\alpha)_i = w_i^\alpha$ if $i > t$. Index the elements of A_t from $t+1$ to $t + \#A_t$ and let $n_t = t + \#A_t$. Then let $x_t^*: R^{n_t} \to R^\infty$ be defined by $x_t^*(y) = \sum_{i=1}^{t} y_i e_i + \sum_{\alpha \in A_t} y_\alpha \tilde{w}_t^\alpha$, where e_i is the ith unit vector of R^∞.

Now let $X_\alpha^t = R^{n_t}$. For each $\alpha \in A_t$, let $w^{\alpha t} = x_t^{*-1}(w^\alpha)$ and define a preference relation $(>)_\alpha^t$ on X_α^t by $y (>)_\alpha y'$ if and only if $x_t^*(y) (>)_\alpha x_t^*(y')$. Then let $\mathscr{E}_t$ represent the economy $(X_\alpha^t, (>)_\alpha^t, w^{\alpha t})_{\alpha \in A_t}$. A price system for economy $\mathscr{E}_t$ will be an element $p \in R^{n_t}$. In what follows, the notation will be considerably simplified if the same symbol x^t is used to denote both $y \in R^{n_t}$ and $x_t^*(y)$. Similarly the same symbol p^t will be used to denote an element $p^t \in R^{n_t}$ and an element $p \in R^\infty$, where $p_i = p_i^t$ for all $i \leqslant t$.[10]

LEMMA 1. *For each subeconomy $\mathscr{E}_t$, there is a* C.E. $(p^t, (x^{\alpha t}))$ *with $p^t > 0$ and $p^t w^{\alpha t} > 0$ for all $\alpha \in A_t$.*

Proof. By construction $\mathscr{E}_t$ is irreducible and $w^t \gg 0$. Then given Assumptions I and II, one may verify that the assumptions of Theorem 3 in McKenzie [7] are satisfied. The Lemma then follows from this theorem and Assumption III. Q.E.D.

Throughout the remainder of the paper, the sequences $(p^t, (x^{\alpha t}))$ will be understood to refer to a C.E. $(p^t, x^{\alpha t})$ for the subeconomy $\mathscr{E}_t$. Also, $x^t \to x$ will mean convergence in the product topology.

LEMMA 2. *There is a subsequence $(\mathscr{E}_t)$ such that for each α, $x^{\alpha t} \to \bar{x}^\alpha$ with $\sum_{\alpha \in A} \bar{x}^\alpha = w$.*

---

[10] $p^t x^t$ will always refer to the inner product of elements in R^{n_t}.

Proof. Since $(p^t, (x^{\alpha t}))$ is a C.E. for each $\mathscr{E}_t$, it follows that $x^{\alpha t} \geqslant 0$ for all $\alpha \in A$ and $\sum_{\alpha \in A_t} x^{\alpha t} = \sum_{\alpha \in A_t} w^{\alpha t} \leqslant w$. It follows from Tychonoff's theorem *[6, p. 32], therefore, that for some subsequence $(\mathscr{E}_t)$, $x^{\alpha t} \to \bar{x}^\alpha \geqslant 0$ for all $\alpha \in A$ and since $\sum_{\alpha \in A_t} w^{\alpha t} \to w$, $\sum_{\alpha \in A} \bar{x}^\alpha = w$.* Q.E.D.

Henceforth, we will restrict attention to the subsequence of economies described in Lemma 2.

LEMMA 3. *For any $\beta, \gamma \in A$, there is a $K > 0$ such that $0 < p^t w^{\gamma t} < K p^t w^{\beta t}$ whenever $\beta, \gamma \in A_t$.*

Proof. Suppose not. Then $\liminf_{t \to \infty} (p^t w^{\beta t}/p^t w^{\gamma t}) = 0$ for some $\beta, \gamma \in A$. Let $B_2 = \{\alpha \in A\colon \liminf_{t \to \infty} p_t w^{\alpha t}/p_t w^{\gamma t} = 0\}$ and $B_1 = A/B_2$. Using the Cantor diagonalization process, we may select a subsequence $(\mathscr{E}_t)$ such that $p^t w^{\alpha t}/p^t w^{\gamma t} \to 0$ for all $\alpha \in B_2$ and $p^t w^{\alpha t}/p^t w^{\gamma t} > k_\alpha > 0$ for all $\alpha \in B_1$. Since $\gamma \in B_1$ and $\beta \in B_2$, both B_1 and B_2 are non-empty. Therefore, by Assumptions III and VI, there is an $\alpha' \in B_1$ such that $\sum_{\alpha \in B_2} w^\alpha + \bar{x}^{\alpha'} (>)_{\alpha'} \bar{x}^{\alpha'}$. Assumption II(i) then implies that $\sum_{\alpha \in B'} w^\alpha + \bar{x}^{\alpha'} (>) \bar{x}^{\alpha'}$ for some finite $B' \subset B_2$. Therefore, again using Assumption II(i), we may choose $\lambda < 1$ such that $\lambda(\sum_{\alpha \in B'} w^\alpha + x^{\alpha' t}) (>)_{\alpha'} x^{\alpha' t}$ for large t. Therefore, since $p^t w^{\alpha t}/p^t w^{\gamma t} \to 0$ for all $\alpha \in B'$ and $p^t w^{\alpha' t}/p^t w^{\gamma t} > k_{\alpha'} > 0$ for all t, we ave $p^t \lambda(\sum_{\alpha \in B'} w^\alpha + x^{\alpha' t}) < p^t w^{\alpha' t}$ for t large. But this implies that $(p^t, x^{\alpha t})$ s not a competitive equilibrium for $\mathscr{E}_t$. Q.E.D.

For the remainder of the proof, we will choose an $\bar{\alpha} \in A_1$ and normalize rices (p^t) by requiring $p^t)$ by requiring $p^t w^{\bar{\alpha} t} = 1$. For any element $x \in R^\infty$, t $\hat{x}_s$ be the element $(\hat{x}_s)_i = x_i$ if $i \leqslant s$, $(\hat{x}_s)_i = 0$ for $i > s$.

LEMMA 4. *There exists a subsequence $(\mathscr{E}_t)$ such that $p^t \to \bar{p} > 0$ with $w^\alpha < K_\alpha < \infty$ for each $\alpha \in A$.*

Proof. For any $i \in N$, choose α_i so that $w_i^{\alpha_i} > 0$. Then Lemma 3 implies e existence of $K_i > 0$ such that $p^t w^{\alpha_i t} < K_i$ for all t, and hence $< K_i/w_i^{\alpha_i} < \infty$ for all t. Therefore, from Tychonoff's theorem, we may elect a subsequence $(\mathscr{E}_t)$ such that $p^t \to \bar{p}$.

If $\bar{p} w^\alpha = \infty$ for some α, then for any $K > 0$, there is an s such that $\cdot \hat{w}_s > K$. But, for $t > s$, $p^t w^{\alpha t} \geqslant \hat{p}_s^t \hat{w}_s^\alpha$. Therefore, since $\hat{p}_s^t \hat{w}_s^\alpha \to \hat{\bar{p}}_s \hat{w}_s^\alpha$, it llows that $p^t w^{\alpha t} > K$ for large t, which contradicts Lemma 3.

Suppose $\bar{p} = 0$. Then consider agent $\bar{\alpha}$. From Assumptions III and IV, it llows that $w + \bar{x}^{\bar{\alpha}} (>)_{\bar{\alpha}} \bar{x}^{\bar{\alpha}}$. Then since $\hat{w}_s \to w$ and $\hat{\bar{x}}_s^{\bar{\alpha}} \to \bar{x}^{\bar{\alpha}}$, it follows from ssumption II(i) that $\hat{w}_s + \hat{\bar{x}}_s^{\bar{\alpha}} (>)_{\bar{\alpha}} x^{\bar{\alpha} t}$ for s and t large. But if $p^t \to 0$, then $(\hat{w}_s + \hat{\bar{x}}_s^{\bar{\alpha}}) < 1$ for t large. Consequently, $(p^t, (x^{\alpha t}))$ is not a C.E. for onomy $\mathscr{E}_t$. Q.E.D.

Proof of Theorem 1. (a) We show first that $\bar{p}w^{\alpha} > 0$. By Assumptions III and VI, $w^{\alpha} + \bar{x}^{\beta} (>)_{\beta}\, \bar{x}^{\beta}$ for some $\beta \in A$. Therefore, by Assumption II(i) $\lambda(\hat{w}_s^{\alpha} + x^{\beta t})(>)_{\beta}\, x^{\beta t}$ for some s, some $\lambda < 1$ and all t sufficiently large. Therefore, $(p^t, (x^{\alpha t}))$ a C.E. implies $p^t \lambda(\hat{w}_s^{\alpha} + x^{\beta t}) > p^t x^{\beta t}$, or $p^t \hat{w}_s^{\alpha} > (1-\lambda)\, p^t x^{\beta t}$. But Lemma 3 implies $p^t x^{\beta t} > k_{\beta} > 0$ for all t. Therefore $p^t \hat{w}_s^{\alpha} > (1-\lambda)\, k_{\beta}$ for t large and hence $\bar{p}\hat{w}_s^{\alpha} > 0$ for some s. This implies $\bar{p}w^{\alpha} > 0$.

(b) We show next that $\bar{p}(w^{\alpha} - \bar{x}^{\alpha}) \leqslant 0$. Suppose not. Then $\bar{p}(\hat{w}_s^{\alpha} - \hat{\bar{x}}_s^{\alpha}) > 0$ for large s, and hence $p^t(\hat{w}_s^{\alpha} - \hat{x}_s^{\alpha t}) > 0$ for large $t > s$. Now Assumptions III and VI imply that $\bar{x}^{\alpha} + w (>)_{\alpha}\, \bar{x}^{\alpha}$. Then from Assumption II(i), it follows that $\hat{x}_s^{\alpha t} + \hat{w}_s (>)_{\alpha}\, \hat{x}_s^{\alpha t}$, for s and t large, and hence by Assumption II(ii), $\hat{x}_s^{\alpha t} + \lambda \hat{w}_s (>)_{\alpha}\, \hat{x}_s^{\alpha t}$ for $0 < \lambda \leqslant 1$. Therefore, since $p^t(\hat{w}_s^{\alpha} - \hat{x}_s^{\alpha t} - \lambda \hat{w}_s) > 0$ for λ near 0 and large t, $(p, (x^{\alpha t}))$ is not C.E. for economy $\mathscr{E}_t$.

(c) Next we show that $\bar{p}\bar{x}^{\alpha} < \infty$. Suppose not. Then for any $K > 0$, there is an s such that $\bar{p}\hat{\bar{x}}_s^{\alpha} > K$, and hence $p^t \hat{x}_s^{\alpha t} > K$ for large $t > s$. But since $p^t \hat{x}_s^{\alpha t} \leqslant p^t x^{\alpha t} \leqslant p^t w^{\alpha t}$, this contradicts Lemma 3.

Paragraphs (a)–(c) establish condition (i).

(d) *To establish condition (ii) suppose* $z (>)_{\alpha}\, \bar{x}^{\alpha}$. *Then from* Assumption II(i), it follows that $\hat{z}_s (>)_{\alpha}\, x^{\alpha t}$ for s and t sufficiently large. Since $(p^t, (x^{\alpha t}))$ is a C.E., we then have $p^t \hat{z}_s > p^t x^{\alpha t} \geqslant p^t \hat{x}_s^{\alpha t}$ for large t and s. Therefore $p^t \to \bar{p}$ implies $\bar{p}\hat{z}_s \geqslant \bar{p}\hat{\bar{x}}_s^{\alpha}$ and hence, by definition, $\bar{p}z \geqslant \bar{p}\bar{x}^{\alpha}$. Now suppose $\bar{p}z = \bar{p}\bar{x}^{\alpha}$. Then from Assumption II(i), $\lambda z (>)_{\alpha}\, \bar{x}^{\alpha}$ for some $\lambda < 1$. But paragraph (c) then implies $p\lambda z < p\bar{x}^{\alpha}$, a contradiction. Therefore, $\bar{p}z > \bar{p}\bar{x}^{\alpha}$ if $z (>)_{\alpha}\, \bar{x}^{\alpha}$. This combined with condition (i) establishes (ii).

(e) Finally, to establish (iii), we note that if $w_i^{\alpha} = 0$ for all but a finite number of i, then $w^{\alpha} = \hat{w}_s^{\alpha}$ for large s. Then for any $t > 0$, Assumptions II an VI imply $p_{\alpha}^t = 0$ for all $t > s$ (since $\bar{w}_{\alpha}^t = 0$). Therefore, since $w_{\beta}^{\alpha t} = 0$ for $\beta \neq \alpha$, we have $0 \leqslant p^t(w^{\alpha t} - x^{\alpha t}) \leqslant p^t(\hat{w}_s^{\alpha} - \hat{x}_s^{\alpha t})$ for all $t > s$. Therefore, $\bar{p}(\hat{w}_s^{\alpha} - \hat{\bar{x}}_s^{\alpha}) \geqslant 0$ and hence $\bar{p}(w^{\alpha} - \bar{x}^{\alpha}) \geqslant 0$. Combined with condition (i) this implies $\bar{p}(w^{\alpha} - \bar{x}^{\alpha}) = 0$. Q.E.D.

5. PARETO OPTIMALITY

It is well known that the competitive equilibrium in models of overlapping generations need not be Pareto optimal. On the other hand, the equilibria established by Bewley in economies with a finite number of infinity lived agents are always Pareto optimal. It may not be immediately obvious, therefore, what the welfare properties are of an equilibrium in an economy with an infinite number of infinitely-lived agents or in economies which combine infinitely-lived agents with finite-lived agents.

In this section, I present a sufficient condition for the competitive equilibrium to be Pareto optimal. If the aggregate endowment of a finite number of agents accounts for a significant fraction of the economy wide endowment in all but possibly a finite number of goods, and if the preferences of these agents are continuous in the product topology and satisfy an appropriate non-satiation condition, then the competitive equilibrium is Pareto optimal.

We will assume throughout this section that A is a countable set and $w^\alpha \in X_\alpha = R_+^\infty$ for all α (i.e., Assumptions I and IV). In order to establish our welfare result, we need first to establish an important restriction on the set of prices which can support competitive equilibrium.

LEMMA 5. *If* $P_\alpha(x)$ *is open relative to* X_α *and non-empty, then* $p \geqslant 0$ *and* $x \in D_\alpha(p)$ *implies* $px \leqslant pw^\alpha < \infty$.[11]

Proof. Suppose not. Then $pw^\alpha = \infty$. Now let $z \in P_\alpha(x)$. Then since $P_\alpha(x)$ is open, $\hat{z}_s \in P_\alpha(x)$ for s large. But $p\hat{z}_s < \infty$. Therefore, $p(w^\alpha - \hat{z}_s) > 0$, which implies $x \notin D_\alpha(p)$. Q.E.D.

It then follows immediately that

THEOREM 3. *If* P_α *is open valued relative to* X_α *for all* $\alpha \in A$ *and the economy* $\mathscr{E}$ *is irreducible, then* $pw^\alpha < \infty$ *for all* $\alpha \in A$.

In particular, Assumptions I to VI imply that any competitive equilibrium price sequence must assign a finite value to the endowment of each agent.

For the remainder of this section, it will be necessary to modify Assumption II to require that preferences be transitive, asymmetric and complete.

ASSUMPTION II′. *For each agent* $\alpha \in A$ *and* $x \in X_\alpha$:

(i) $(\geqslant)$ *is a transitive, asymmetric, and complete preference ordering and is continuous in the product topology.*

(ii) *If* $z (>)_\alpha x$, *then* $\lambda z + (1 - \lambda) x (>)_\alpha x$ *for* $0 < \lambda \leqslant 1$.

Then, as usual, we will say that an attainable allocation (x^α) is *Pareto optimal* if there is no other attainable allocation (z^α) such that $z^\alpha (\geqslant)_\alpha x^\alpha$ for all $\alpha \in A$ and with $z^\alpha (>)_\alpha x^\alpha$ for some $\alpha \in A$.

The main result of this section is based on the following proposition, which is little more than a restatement of a well-known result by Debreu [5].

THEOREM 4. *Suppose* $(p, (x^\alpha))$ *with* $p \geqslant 0$ *is a C.E. for economy* $\mathscr{E}$. *Then if* $\mathscr{E}$ *is irreducible and Assumption II′ is satisfied,* $pw < \infty$ *implies* (x^α) *is Pareto optimal.*

---

[11] If $p \not\geqslant 0$, then the corresponding Lemma requires $\sum_{i=1}^{\infty} |p_i w_i^\alpha| < \infty$.

Proof. We show first that $z (\succcurlyeq)_\alpha x^\alpha$ and $z \leqslant w$ implies $p(w^\alpha - z) \leqslant 0$. Suppose not. Then since $\mathscr{E}$ is irreducible, there is a $y \leqslant w$ such that $z + y (>)_\alpha z$. Therefore, by Assumption II′, $z + \lambda y (>)_\alpha x^\alpha$ for $0 < \lambda \leqslant 1$. But if $pz < pw^\alpha$, and $pw < \infty$, $y \leqslant w$ implies $p(z + \lambda y) < pw^\alpha$ for some λ near 0. Hence $x^\alpha \notin D_\alpha(p)$.

Now let (z^α) be an attainable allocation with $z^\alpha (\succcurlyeq)_\alpha x^\alpha$ for all α and $z^\beta (>)_\beta x^\beta$. Then $p(w^\alpha - z^\alpha) \leqslant 0$ for all α and $p(w^\beta - z^\beta) < 0$. But $pw < \infty$ then implies $\sum_{\alpha \in A} p(w^\alpha - z^\alpha) = pw - p \sum_{\alpha \in A} z^\alpha < 0$, which contradicts $w = \sum_{\alpha \in A} z^\alpha$. Q.E.D.

Note that, as usual, the continuity of preferences is not needed for this proof.

Combining Lemma 5 with Theorem 4 we then obtain

THEOREM 5. *Suppose Assumption II′ is satisfied and the economy $\mathscr{E}$ is irreducible. Then if there is a finite set of agents $B \subset A$ and an $\varepsilon > 0$ such that $\sum_{\alpha \in B} w_i^\alpha > \varepsilon w_i$ for all but a finite number of i, (x^α) is Pareto optimal for any C.E. $(p, (x^\alpha))$.*

Proof. From Lemma 5, $pw^\alpha < \infty$ for all $\alpha \in A$. Then there is a $K < \infty$ such that $pw \leqslant (1/\varepsilon) p \sum_{\alpha \in B} w^\alpha + K = (1/\varepsilon) \sum_{\alpha \in B} pw^\alpha + K < \infty$. The theorem then follows from Theorem 3. Q.E.D.

As is clear from the proof of Theorem 5, the importance of assuming a finite set of agents with a significant fraction of the aggregate endowment is to guarantee the validity of Walras' Law (i.e., $pw = p \sum x^\alpha < \infty$). If this condition is not satisfied, then the equilibrium value of the aggregate endowment may become infinite and the resulting allocation may not be Pareto optimal. This is not to say that the competitive allocation is not Pareto optimal unless the value of the endowment is finite. It is easy to construct examples where that is not the case. However, a complete characterization of the price sequences which support Pareto optima is beyond the scope of this paper.[12]

6. ALTERNATIVE ASSUMPTIONS

In this section, I want to examine more closely the role of some of the assumptions used to establish Theorems 1 and 2 and in some cases indicate how they might be modified.

---

[12] Significant progress in this direction for economies with overlapping generations has been made by Bose [4] and more recently by Balasko and Shell [1].

6.1. *Irreducibility*

The need for a resource relatedness assumption is already well understood from the general competitive theory with only a finite number of goods and agents. Nevertheless, it may be useful to present a simple example in the context of dynamic model with overlapping generation which illustrates the problems which may arise when the economy is not irreducible.

Suppose the agents are indexed by the natural numbers, N. Let the endowment of agent α be given by $w_\alpha^\alpha = 1$, $w_i^\alpha = 0$ if $i \neq \alpha$ and his preferences be represented by the utility function $u^\alpha(x) = x_\alpha \cdot x_{\alpha+1}$. This economy is not irreducible. (To see this, let $B_1 = \{1\}$, $B_2 = N/\{1\}$.) There is also no competitive equilibrium. Suppose p were an equilibrium price sequence. Then if $p_1 > 0$, agent 1 would choose to sell $\frac{1}{2}$ of his endowment of good 1. But no other agent would be willing to buy any of the good at a positive price. Therefore $p_1 = 0$. Similarly, $p_2 > 0$ implies that agent 2 would sell $\frac{1}{2}$ of his endowment of good 2. But since $p_1 = 0$ implies $pw^1 = 0$, no other agent would buy the good at a positive price. Therefore, $p_2 = 0$. But, if $p_2 = 0$, then agent 1 would demand an infinite amount of the good, which is inconsistent with market clearing. Consequently, there is no equilibrium.

6.2. *Free Disposal*

As I have noted already in several places, some of the assumptions used to establish Theorems 1 and 2 could have been relaxed in a number of respects without changing the basic results. In general, the more restrictive assumptions were used in order to simplify notation and make the central idea behind the proofs more transparent. In a more general model, particularly one including production, it might be desirable to explicitly relax some of these assumptions. One such example is the assumption of free disposal (Assumption III). In fact, many of the complications which appear when this assumption is relaxed appear anyway when production is introduced.

Since free disposal is not needed to establish the existence of an equilibrium in a finite good economy (see McKenzie [7]), the only problem is in the extension to the infinite-good, infinite-agent economy. Here two problems emerge. First, since prices may be negative, we must distinguish between convergence and absolute convergence, i.e., between $\sum_{i=1}^{\infty} p_i x_i$ and $\sum_{i=1}^{\infty} |p_i x_i|$. For example, to establish the results in Theorem 1, it is not sufficient merely to show that $pw^\alpha < \infty$; in fact, what we require is that $\sum_{i=1}^{\infty} |p_i w_i^\alpha| < \infty$. Fortunately, this result can also be demonstrated. The main difference in the proof is that rather than normalizing p^t by $p^t w^{\bar{\alpha}} = 1$ as was done in this paper, it is necessary instead that prices be normalized by $\sum_{i=1}^{\infty} |p_i^t w_i^{\bar{\alpha}}| = 1$.

The second problem stems from the fact that *it* may no longer be appropriate to assume $X_a = R_+^\infty$ if we are to retain the necessary continuity restrictions on preferences. This in turn requires that the irreducibility condition (Assumption VI) be reformulated with some care in order to guarantee the existence of equilibrium in each of the finite agent subeconomies considered in Lemma 1. Essentially, we need to ensure that each of the subeconomies are irreducible when the commodity spaces are restricted to some appropriately chosen finite-dimensional subspace of R^∞.

6.3. *Mackey Continuity*

One difference in the approach used by Bewley [3] and the approach used here is in the choice of the commodity space and the corresponding topology in which the continuity of preferences are assumed. Since Bewley allowed the set of commodities to be represented by an arbitrary sigma-finite measure space, it was necessary that he restrict the commodity space to the set of all bounded measurable functions, L_∞. He was then able to establish his results under the assumption that preferences were Mackey continuous. Since I consider only discrete economies (i.e., with only a countable number of goods), I chose instead to allow the commodity space to be represented by the set of all non-negative sequences and require that preferences be continuous in the product topology. I should emphasize, however, that the argument does not change at all if, following Bewley, I were to identify the commodity space of each agent with l_∞^+ and assume that preferences are Mackey continuous. In this case, however, it is essential that we first normalize goods so that the aggregate endowment also lies in l_∞; otherwise, there may be no equilibrium. In particular, it is not sufficient to assume merely that the endowment of each individual agent is bounded. This can be interpreted as a requirement that each agent discount future consumption at a rate which on average exceeds not only the average growth rate of his own endowment but the average growth rate of aggregate endowment as well. The following example illustrates the problems which emerge if this condition is not satisfied.

Suppose the set of agents are indexed by $\{0, 1, 2,...,\}$. For $\alpha \geqslant 1$, let $w_i^\alpha = 2^i$ for $i = \alpha, \alpha + 1$ and 0 otherwise and let $u^\alpha(x) = x_\alpha + \gamma x_{\alpha+1}$, where $\gamma < \frac{1}{2}$. For $\alpha = 0$, let $w^0 = (2, 0, 0,...,)$ and for $x \in l_+^\infty$ let $u^0(x) = \sum_{i=1}^\infty x_i \beta^i$, where $1 > \beta > \frac{1}{2}$. Then the preferences of each agent are Mackey continuous and the endowment of each agent is bounded. However, the aggregate endowment is equal to (4, 8, 16,...,) which is not bounded.

Now suppose p is the equilibrium price vector. Note first that we must have $p_{i+1} \leqslant \frac{1}{2} p_i$ for all i. Otherwise utility maximization implies $x_{i+1}^i = 0$, and hence, by the budget constraint, $p_i x_i^i = p_{i+1} 2^{i+1} + p_i 2^i$, which implies (since $p_{i+1} > \frac{1}{2} p_i$), $x_i^i > 2^{i+1} = w_i$. Consequently, the market would not clear. But if $p_{i+1} \leqslant \frac{1}{2} p_i$ for all i, then since $\beta \geqslant \frac{1}{2}$, there is no most preferred element in the budget set of agent 0. Hence, a competitive equilibrium (with or without transfer payments) does not exist.

7. AN EXAMPLE OF EQUILIBRIUM ONLY WITH POSITIVE TRANSFER PAYMENTS

The results of Section 2 only imply the existence of a competitive equilibrium when either (1) the endowment of some finite set of agents is large relative to the aggregate endowment, or (2) each agent has a positive endowment for only a finite number of goods. Otherwise, the most that is guaranteed is an equilibrium with non-negative transfer payments. In this section, I will provide an example of an economy where there is an equilibrium with transfer payments but no competitive equilibrium. This may also be interpreted as an example where only Samuelson-type monetary equilibria exist.[13]

Index the set of agents by $\{0, 1, 2, \ldots, \}$. For $\alpha \geqslant 1$, let $w_i^\alpha = 1$ if $i = \alpha, \alpha + 1$ and 0 otherwise and let $u^\alpha(x) = x_\alpha + 3x_{\alpha+1}$. For agent 0, let $w_i^0 = \frac{1}{2}^i$ for all i and $u^0(x) = x_1$.

Note first that in the absence of agent 0, the competitive equilibrium price vector must satisfy $p_i = 3p_{i-1}$ for all i with each agent consuming his own endowment. When agent 0 is added, however, this can no longer be the equilibrium price, since the value of agent 0's endowment will be infinite and all of his demand will directed toward first period consumption. We will show first that there are in fact no competitive equilibrium prices.

Let $I = \{i > 1\colon\ p_{i+1} < 3p_i\}$. Then for any $i \in I$, utility maximization implies $x_i^i = 0$. Therefore, $x_i^{i-1} = w_i^i + w_i^{i-1} + w_i^0 > 2$. The budget constraint of agent $i-1$ then implies $2p_i < p_i x_i^{i-1} = p_{i-1} w_{i-1}^{i-1} + p_i w_i^{i-1} = p_{i-1} + p_i$. From this it follows that $i \in I$ implies $p_{i-1} > p_i$. In particular, $i \in I$ implies $i - 1 \in I$. But the demand of agent 0 will only be well defined if $pw^0 < \infty$. Therefore, I must be an infinite set. It then follows from induction that $i \in I$ for all $i > 1$.

This result then implies that for all i, $p_{i+1} < p_i$ and hence, $p_i < p_1$ for all $i > 1$. Therefore, $pw^0 = \sum_{i=1}^{\infty} p_i 2^{-i} < p_1$, which from the budget constraint of agent 0 implies $x_1^0 < 1$. However, $p_2 < p_1$ also implies that $x_1^1 = 0$. But since $w_1^0 + w_1^1 = \frac{3}{2} > 1 > x_1^0 + x_1^1$, it follows that a competitive equilibrium cannot exist.

---

[13] The example in Section 6.1 is also a case where only monetary equilibria exist. However, competitive equilibria does not exist for any truncated finite-agent subeconomy either. Hence, the problem with the existence of equilibrium in that example has nothing to do with an infinity of agents.

This example is particularly instructive because it illustrates that it is the nature of his endowment and not his preferences which define the essential features of a finite-lived agent, at least for guaranteeing the existence of a competitive equilibrium. Even if an agent consumes only one good, as long as his endowment is positive for an infinite number of goods, he is essentially infinitely lived. On the other hand, as Theorem 2 demonstrates, it does not matter whether agents spread their consumption over a finite or infinite number of goods, as long as his endowment of each agent is positive for only a finite number of goods, an equilibrium must exist.

Although there is no competitive equilibrium in this example, it is not difficult to find a price sequence which generates an equilibrium with non-negative transfer payments. For each $i \geqslant 1$, let $p_i = p_{i-1}/(1 + 2^{-i})$. Then define the allocation (x^α) by $x_i^\alpha = 2 + 2^{-i}$ for $i = \alpha + 1$ and 0 otherwise if $\alpha \geqslant 1$ and $x^0 = (\frac{3}{2}, 0, 0, \ldots,)$. Then (x^α) is feasible, $x^\alpha \in D_\alpha(p)$ for $\alpha \geqslant 1$, and $u^0(x^0) \geqslant u^0(x)$ for any $x \in R_+^\infty$ for which $px \leqslant px^0$. Therefore, $(p, (x^\alpha))$ is a competitive equilibrium with non-negative transfer payments. It is not a competitive equilibrium, however. Since $pw^0 < p_1$ and $px^0 = \frac{3}{2} p_1$, agent 0 must be awarded a positive transfer payment of more than $\frac{1}{2} p_1$ if he is to be able to purchase x^0.

As was the case in our discussion of Pareto optimality, the possibility that positive transfer payments may be needed to clear all the markets may be traced back to the absence of Walras Law. When the value of the aggregate endowment is infinite, it is possible to construct an attainable allocation such that the value of each agent's consumption exceeds the value of his endowment.

As I noted at the outset, an equilibrium with positive transfer payments can be interpreted as an equilibrium with Samuelson's money. Suppose, in the example just presented, we were to augment with endowment of agent 0 with an amount of Samuelson's money so that he could just purchase $\frac{3}{2}$ units of good 1 at the equilibrium prices. Then the equilibrium allocation could be attained through the following sequence of transactions. Agent 0 purchases agent 1's endowment of good 1 with money and promises to deliver agent 0's endowment of future goods. Agent 1 then purchases agent 2's endowment of good 2 with the money obtained from agent 0 plus promises to deliver the remainder of agent 0's endowment. This process continues forever. Given this interpretation, what we have constructed is an example of an economy where, not only do monetary equilibria exist, but where there can be no equilibrium without a positive value of money, even though all the conditions necessary to guarantee the existence of equilibrium in a finite good economy are satisfied.

8. CONCLUSION

This paper makes three basic points. First, it illustrates, through an example, that in a dynamic model with an infinity of agents, a competitive equilibrium need not exist even when all the conditions sufficient for the classical existence theorems are satisfied. The same example also illustrates how Samuelson's money might be introduced to generate a monetary equilibrium in those cases when the competitive equilibrium does not exist.

Second, it provides a method for establishing the existence of a competitive equilibrium in those cases where one does exist. I have indicated in several places how the proof should proceed if some of the assumptions used here were modified. It should also be possible to incorporate production. Furthermore, I believe that the basic idea can be extended to cover economies which contain a continuum of goods, as in Bewley [3].

Finally, the paper provides a sufficient condition for competitive equilibria to be Pareto optimal. Although considerably more work needs to be done in this area, the results here do emphasize an important relationship between the number of agents and the validity of Walras Law.

REFERENCES

1. Y. BALASKO AND K. SHELL, The overlapping-generations model, I. The case of pure exchange without money, *J. Econ. Theory* **23** (1980), 281–306.
2. Y. BALASKO, D. CASS, AND K. SHELL, Existence of competitive equilibrium in a general overlapping-generations model, *J. Econ. Theory* **23** (1980), 307–322.
3. T. BEWLEY, Éxistence of equilibrium in markets with many commodities, *J. Econ. Theory* **4** (1972), 514–540.
4. A. BOSE, "Pareto Optimality and Efficient Capital Accumulation," University of Rochester Discussion Paper 74–4, (1974).
5. G. DEBREU, Valuation equilibria and pareto optima, *Proc. Nat. Acad. Sci. U.S.A.* **40** (1954), 588–592.
6. N. DUNFORD AND J. SCHWARTZ, "Linear Operators, Part I" Interscience, New York, 1957.
7. L. W. MCKENZIE, The classical existence theorem of competitive equilibrium, *Econometrica*, in press.
8 P. A. SAMUELSON, An exact consumption-loan model of interest with or without the social contrivance of money, *J. Pol. Econ.* **66** (1958), 467–482.

JOURNAL OF ECONOMIC THEORY 24, 112–142 (1981)

The Overlapping-Generations Model. II. The Case of Pure Exchange with Money*

YVES BALASKO

Université Paris I, CEPREMAP, Paris, France, and
Center for Analytic Research in Economics and the Social Sciences,
University of Pennsylvania, Philadelphia, Pennsylvania 19104

AND

KARL SHELL

Center for Analytic Research in Economics and the Social Sciences,
University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received December 14, 1979; revised February 25, 1980

1. INTRODUCTION

Government debt instruments (e.g., money and government bonds) serve many functions in the private sector. They can be stores of value, vehicles for the payment of taxes, media for transactions, and so forth. Their roles vary from economy to economy depending upon institutions and conventions. One of the roles, however, is basic to all others: if a government debt instrument does not serve as a value store, then it cannot serve any other useful function.

In the present article, the paper assets created by the government are assumed to have two functions: (1) they are, of course, potential stores of value; and (2) they are used by households in paying their taxes and by the government in making transfers. In addition, we make the extreme assumptions that markets are complete and transactions are costless. There is then no essential economic difference between the various forms of government debt; hence, we shall refer to all forms of "government paper" as simply "money."

Money does not in general serve as a proper store of value—i.e., money cannot have a positive price—in the finite-horizon economy in which the

---

* This research was supported by Grants SOC 78-06157 and SES 80-07012 from the National Science Foundation to the Center for Analytic Research in Economics and the Social Sciences at the University of Pennsylvania.

0022-0531/81/010112-31$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

terminal date is known with certainty. The reason is obvious. Money is worthless at the end of the final period. Consequently, in the next-to-last period, individuals desire to dispose of money holdings in order to avoid capital losses. This drives the price of money to zero at the end of the next-to-last period. And so on. Individuals with foresight drive the price of money to zero in each period, i.e., the "general price level" in equilibrium must be infinite.

The natural way to permit money to be a proper store of value is to go beyond the finite-horizon model. Our analysis is cast, therefore, in terms of the overlapping-generations economy, first analyzed by Samuelson [17]. Samuelson's model is especially attractive for macroeconomic analysis. It is genuinely dynamic, reflecting the vitality and mortality of consumers, and the unbounded, unidirectional flow of time. Furthermore, the model is basically disaggregative—clearly distinguishing between the tastes and opportunities of the individual consumers.

In the present paper, we treat a pure-exchange economy. We go well beyond the analysis of [9, 11, 13, 15, 17, 18], allowing for many commodities in each period, for tastes and endowments which vary from individual to individual, and for *arbitrary* government monetary policies. In Section 2, the basic model is introduced and the concept of *monetary competitive equilibrium* is defined. The consequences of equilibrium for the money markets are examined in Section 3. We then define the closely related, but mathematically simpler, concept of *Walrasian equilibrium*. We focus on equilibrium in the markets for real commodities in Section 4.

A *monetary policy* is a sequence of money taxes and money transfers. For every monetary policy, there is at least one competitive equilibrium. Money, however, might be irrelevant in the sense that the equilibrium price of money is identically zero. In Section 5, we study *bonafide monetary policies*, those policies which permit the existence of a nontrivial price of money. A sequence of commodity prices and a sequence of individual incomes is said to be a *price-income equilibrium* if aggregate demands for commodities at those prices and incomes is equal to the given aggregate resources of the economy. We provide a characterization of the monetary policies consistent with a given price-income equilibrium. We proceed to establish that if long-run interest rates exceed long-run growth rates, then bonafide monetary policies entail long-run money stocks of zero. On the other hand, if asymptotically the money supply is growing faster than the real growth rates, then any proper monetary equilibrium is not Pareto-optimal.

In Section 6, we show that the set of bonafide monetary policies is connected. It is also shown that any Pareto-optimal allocation can be achieved as a competitive equilibrium if the government has control over monetary policy and the price of money, even if endowments of real commodities cannot be reassigned. A monetary policy is said to be *potent* if for some price of money the resulting equilibrium allocation is Pareto-optimal. We show that the set of potent monetary policies is connected.

The set of equilibrium money prices, given fixed endowments and a fixed monetary policy, is investigated in Section 7. A natural bijection relates this set to a subset of the set of bonafide monetary policies. If the monetary policy is neither trivial nor pathological, the set of equilibrium money prices is bounded. Typically, there is a vast multiplicity of such monetary equilibria; this has important consequences for monetary theory and policy.

In the Appendix, we establish the connectedness of the set of weakly-Pareto-optimal allocations and the arc-connectedness of the set of Pareto-optimal allocations in the overlapping-generations model. A list of some basic notation appears as a Glossary, which follows the Appendix.

2. THE MONETARY ECONOMY

The pure-exchange, overlapping-generations model of Balasko and Shell [5] is extended to allow for consumer money holdings; the basic assumptions about the "real" economy along with most of the corresponding notation are taken from [5]. Our extension to the monetary case allows the government to pursue any arbitrary (active or inactive) monetary policy; no restriction is placed on the sequence of nominal lump-sum money transfers and money taxes.

We summarize the basic features of the formal model. For analytic convenience, we adopt the simple demographic pattern analysed in [5], but recall that this is done without serious loss in generality (see [4, Sect. IV]). Consumers, who are indexed by their order of birth $h = 0, 1,\dots$ are either present at the inception of the economy (in which case they live out the balance of their lives during period 1) or are born at the beginning of some period t ($t = 1, 2,\dots$) and live out the whole of their lives in periods t and $t+1$. Each generation consists of a single consumer indexed uniquely by its birthdate t ($t = 0, 1,\dots$).

In each period t ($t = 1, 2,\dots$) there are l completely perishable commodities. There is also a completely imperishable fiat money. There is no production. Let $x_t^{s,i}$ be consumption of commodity i ($i = 1,\dots, l$) by consumer t in period s. Consumer t derives utility from consuming goods during his lifetime; thus utility functions can be represented as

$$u_t(x_t) \qquad \text{for} \quad t = 0, 1,\dots,$$

where

$$x_0 = x_0^1 = (x_0^{1,1},\dots, x_0^{1,l}) \in \mathbb{R}^l_{++} \qquad \text{for} \quad t = 0$$

and

$$x_t = (x_t^t, x_t^{t+1}) = (x_t^{t,1},\dots, x_t^{t,l}, x_t^{t+1,1},\dots, x_t^{t+1,l}) \in \mathbb{R}^{2l}_{++} \qquad \text{for} \quad t = 1, 2,\dots .$$

When convenient, we also denote by x_0 and x_t the respective sequences

$$x_0 = (x_0^1, 0,\dots, 0,\dots)$$

and

$$x_t = (0,\dots, 0, x_t^t, x_t^{t+1}, 0,\dots) \qquad \text{for} \quad t \geqslant 1.$$

Let $x = (x_0, x_1,\dots)$ be the sequence of commodity allocations and X be the set of all such allocation sequences, i.e., $X = \mathbb{R}^l_{++} \times \mathbb{R}^{2l}_{++} \times \mathbb{R}^{2l}_{++} \times \cdots$.

The utility functions $u_0(\cdot)$ and $u_t(\cdot)$ are assumed to have strictly positive first-order partial derivatives (i.e., to be smoothly monotonic) and to be strictly quasi-concave (i.e., to exhibit diminishing marginal rates of commodity substitution). Furthermore, in order to rule out complicated boundary behavior, the closure of every indifference surface in $\mathbb{R}^l$ (resp. $\mathbb{R}^{2l}$) is assumed to be contained in the corresponding strictly positive orthant $\mathbb{R}^l_{++}$ (resp. $\mathbb{R}^{2l}_{++}$). Each consumer has strictly positive endowments of the commodities available during his lifetime,

$$\omega_0 = \omega_0^1 = (\omega_0^{1,1},\dots, \omega_0^{1,l}) \in \mathbb{R}^l_{++}$$

and

$$\omega_t = (\omega_t^t, \omega_t^{t+1}) = (\omega_t^{t,1},\dots, \omega_t^{t,l}, \omega_t^{t+1,1},\dots, \omega_t^{t+1,l}) \in \mathbb{R}^{2l}_{++} \qquad \text{for } t \geqslant 1.$$

When convenient, we also denote by ω_0 and ω_t the respective sequences

$$\omega_0 = (\omega_0^1, 0,\dots, 0,\dots)$$

and

$$\omega_t = (0,\dots, 0, \omega_t^t, \omega_t^{t+1}, 0,\dots) \qquad \text{for} \quad t \geqslant 1.$$

Let $\omega = (\omega_0, \omega_1,\dots)$ denote the positive sequence of commodity endowments in X.

We turn to the monetary aspects of the model. The government cannot redistribute commodity endowments but can create nominal money and distribute it to individuals. It can also levy nominal money taxes on individuals. We assume that transfers and taxes are made in a lump-sum fashion during the lifetimes of individual consumers and that they are perfectly foreseen. As a result of the government's policy, therefore, consumers can be considered to be "endowed" with nominal money (in possibly positive, zero, or negative amounts) for each period of their lives. Monetary transfers (alternatively, endowments of money) are

$$m_0^1 \in \mathbb{R} \qquad \text{for} \quad t = 0$$

and

$$(m_t^t, m_t^{t+1}) \in \mathbb{R}^2 \qquad \text{for} \quad t = 1, 2, \ldots,$$

where m_t^s is the lump-sum money transfer to consumer t in period s. Since only the living can receive transfers or pay taxes, we have that for $t \geqslant 1$, $m_t^s = 0$ if $s < t$ or $s > t + 1$, while $m_0^s = 0$ for $s \neq 1$. The nominal supply of money extant at time t is

$$m^t = \sum_{s=1}^{t} \sum_{\tau} m_\tau^s \in \mathbb{R},$$

where $\sum_\tau m_\tau^s$ is the aggregate amount of money created by the government in period s.

The government's control of the economy is exercised through its monetary transfers, which can be summarized by the sequence $m = (m_0^1, m_1^1, m_1^2, \ldots, m_t^t, m_t^{t+1}, \ldots)$ in the space of feasible monetary transfers, $M = \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \times \cdots$. With only one type of government debt instrument, no useful distinction can be drawn between monetary and fiscal policy. We adopt the convention of the overlapping-generations literature by referring to the sequence $m \in M$ as *monetary policy*. The reader should be warned, however, that in the classical macroeconomics literature the sequence $m \in M$ is frequently referred to as fiscal policy, leaving for "monetary policy" the composition of the public debt between monetary and nonmonetary instruments.

Just as an individual chooses his lifetime consumption profile, he must also choose a lifetime profile of money holdings. Let $x_t^{s,m} \in \mathbb{R}$ be the gross addition to his holdings of money committed by consumer t in period s. If $x_t^{s,m}$ is positive, he is increasing his inventory of money; if negative, he is decreasing his inventory. If $(x_t^{s,m} - m_t^s)$ is positive, consumer t is a net purchaser of money in period s, since he is committing to inventory more money than he receives in transfers. If $(x_t^{s,m} - m_t^s)$ is negative, he is a net seller of money. If $x_t^{s,m} = m_t^s$, he is neither a net purchaser nor a net seller of money: he salts away his entire transfer, or, alternatively, he purchases only enough money to cover his taxes.

Only the government is permitted to create outside money, i.e., money which is not redeemed. Hence, all individual money holdings are constrained to be nonnegative at death, yielding constraints which take the form

$$x_0^{1,m} \geqslant 0$$

and (2.1)

$$x_t^{t,m} + x_t^{t+1,m} \geqslant 0 \qquad \text{for} \quad t \geqslant 1,$$

under the realistic convention that $x_t^{s,m} = 0$ for $s > t + 1$.

Define the sequence of gross additions to money holdings by

$$x^m = (x_0^{1,m}, x_1^{1,m}, x_1^{2,m}, \dots, x_t^{t+1,m}, \dots).$$

Each consumer can buy and sell on the spot market and the forward market at perfectly foreseen prices. Let $p^{t,i}$ denote the price of commodity i in period t. Denote by p^t the vector $(p^{t,1}, \dots, p^{t,l}) \in \mathbb{R}^l_{++}$ and by p the price sequence $(p^1, p^2, \dots)$. We choose the normalization $p^{1,1} = 1$ and thus restrict attention to the set of sequences of *present* prices,

$$\mathcal{P} = \{p \mid p^{1,1} = 1\}.$$

Denote by $p^{t,m} \in \mathbb{R}_+$ the *present* price (in terms of the *numéraire* commodity) of money delivered in period t.

Each consumer's lifetime consumption profile and profile of money holdings is then the solution of a budget-constrained utility-maximization problem

$$\begin{aligned}
&\text{maximize } u_0(x_0^1)\\
&\text{subject to}\\
&\qquad p_1 \cdot x_0^1 + p^{1,m} x_0^{1,m} \leqslant p^1 \cdot \omega_0^1 + p^{1,m} m_0^1 = w_0,\\
&\qquad x_0^{1,m} \geqslant 0, \qquad x_0^1 \geqslant 0, \qquad\qquad\qquad \text{for } t = 0,\\
&\text{maximize } u_t(x_t^t, x_t^{t+1}) \qquad\qquad\qquad\qquad\qquad\qquad (2.2)\\
&\text{subject to}\\
&\qquad p^t \cdot x_t^t + p^{t+1} \cdot x_t^{t+1} + p^{t,m} x_t^{t,m} + p^{t+1,m} x_t^{t+1,m}\\
&\qquad \leqslant p^t \cdot \omega_t^t + p^{t+1} \cdot \omega_t^{t+1} + p^{t,m} m_t^t + p^{t+1,m} m_t^{t+1}\\
&\qquad = w_t,\\
&\qquad x_t^{t,m} + x_t^{t+1,m} \geqslant 0, \qquad (x_t^t, x_t^{t+1}) \geqslant 0, \qquad \text{for } t \geqslant 1,
\end{aligned}$$

where w_t then denotes the present value of consumer t's wealth.

2.3 DEFINITION. Let $\omega \in X$ be a sequence of strictly positive commodity endowments and $m \in M$ be a sequence of monetary transfers. A *monetary competitive equilibrium* associated with ω and m is a sequence of commodity prices $p \in \mathscr{P}$ and a sequence of nonnegative money prices $(p^{1,m}, p^{2,m},\ldots)$ (along with the associated consumption profiles $x \in X$ and individual gross monetary additions x^m) satisfying the market-clearing conditions

$$\sum_t x_t = \sum_t \omega_t$$

and

$$x_{t-1}^{t-1,m} + x_{t-1}^{t,m} + x_t^{t,m} \leqslant m^t \qquad \text{for} \quad t \geqslant 1,$$

where $x_0^{0,m} = 0$, and $(x_{t-1}^{t-1,m} + x_{t-1}^{t,m} + x_t^{t,m})$ and m^t are, respectively, the aggregate demand for and supply of money in period t, and where x and x^m are solutions to the system (2.2) for given ω and m.

3. EQUILIBRIUM IN THE MONEY MARKETS

We derive some basic consequences of equilibrium in the money markets. These results are of interest in their own right, and will be used in Section 4 to develop a simplified representation of (Walrasian) equilibrium, which can replace the monetary-competitive-equilibrium concept defined in Secion 2.

3.1 PROPOSITION. *Let the price sequences $p \in \mathscr{P}$ and $(p^{1,m}, p^{2,m},\ldots) \geqslant 0$ be a monetary competitive equilibrium for the endowment sequence $\omega \in X$ and the monetary transfer sequence $m \in M$. Then, the present price of money is a nonnegative constant, i.e., $p^{t,m} = p^m \geqslant 0$ for $t = 1, 2,\ldots$.*

Proof. Assume that for some t, $p^{t,m} \neq p^{t+1,m}$. We must show that this is inconsistent with monetary competitive equilibrium. The budget constraints of consumer t ($t \geqslant 1$) can be rewritten (from system (2.2)) as

$$\begin{aligned} p^t \cdot x_t^t + p^{t+1} \cdot x_t^{t+1} \leqslant p^t \cdot \omega_t^t + p^{t+1} \cdot \omega_t^{t+1} + p^{t,m}(m_t^t - x_t^{t,m}) \\ + p^{t+1,m}(m_t^{t+1} - x_t^{t+1,m}) \end{aligned}$$

$$x_t^{t,m} + x_t^{t+1,m} \geqslant 0.$$

Assume, for example, that $p^{t+1,m} > p^{t,m}$. By setting $x_t^{t+1,m} = -x_t^{t,m}$ and choosing an arbitrarily large purchase of money in period t, so that $x_t^{t,m}$ is arbitrarily large, along with a correspondingly large sale of money in period $t+1$, so that $(-x_t^{t+1,m})$ is arbitarily large, consumer t is able to afford any consumption profile $(x_t^t, x_{t+1}^t) \in \mathbb{R}^{2l}_{++}$. Since the utility function $u_t(\cdot)$ is strictly increasing this entails a contradiction to the equilibrium requirement that materials balance in the commodity market, i.e., $\sum_t x_t = \sum_t \omega_t$. ▮

3.2 *Remark.* Proposition (3.1) is a zero-profit condition, excluding the possibility of arbitrage from intertemporal trading in money.

Since no separate sign restrictions are imposed on $x_t^{t,m}$ or $x_t^{t+1,m}$, the model can be interpreted as having perfect borrowing and lending markets, or perfect markets for outside money *and* inside money. Indeed, the constancy of the *present* price of money p^m has a familiar interpretation equivalent to the standard conditions for perfect "capital" markets. In our model, money serves only as a store of value. Also, the (nominal) own rate of interest on money is assumed to be zero. Hence money's *current* price in terms of the ith commodity, $(p^{t,m}/p^{t,i})$, must in equilibrium increase at the commodity rate of interest $\rho^{t,i}$, where by definition we have

$$\rho^{t,i} = \frac{p^{t,i} - p^{t+1,i}}{p^{t+1,i}} = \frac{p^{t,i}}{p^{t+1,i}} - 1.$$

The current commodity price of money, $(p^{t,m}/p^{t,i})$, is increasing at the commodity interest rate if and only if the present price of money, $p^{t,m}$, is constant $(=p^m \geqslant 0)$.

Proposition (3.2) allows us to employ a more condensed notation. We say that $q \in \mathcal{Q} = \{(p, p^m) \mid p \in \mathcal{P} \text{ and } p^m \in \mathbb{R}_+\}$ defines a monetary competitive equilibrium for $\omega \in X$ and $m \in M$ if the sequences p and $(p^m, p^m, \ldots)$ satisfy Definition (2.3).

3.3. PROPOSITION. *Let the sequence $q = (p, p^m) \in \mathcal{Q}$ define a monetary competitive equilibrium for $\omega \in X$ and $m \in M$. Then, it follows that*

$$p^m x_0^{1,m} = 0$$

and

$$p^m (x_t^{t,m} + x_t^{t+1,m}) = 0 \qquad \textit{for} \quad t \geqslant 1.$$

Proof. Because utility functions are strictly increasing, a solution to the budget-constrained utility maximization system (2.2) must have $x_0^{1,m} = 0$ and $x_t^{t,m} + x_t^{t+1,m} = 0$ whenever $p^m > 0$. ▮

There is no individual bequest motive in our model. Proposition (3.3) states that when money has positive value, it is not optimal for any individual to die holding positive money balances.

3.4. PROPOSITION. *Let $q = (p, p^m) \in \mathscr{Q}$ define a monetary competitive equilibrium for given sequences $\omega \in X$ and $m \in M$. Then the associated equilibrium allocations, $x \in X$ and $x^m \in \mathbb{R}^\infty$, must satisfy the following constraints*:

$$p^1 \cdot x_0^1 = -p^m(x_0^{1,m} - m_0^1) + p^1 \cdot \omega_0^1 \qquad \textit{for} \quad t = 0$$

and

$$\begin{aligned} & p^t \cdot x_t^t + p^m(x_t^{t,m} - m_t^t) = p^t \cdot \omega_t^t, \\ & p^{t+1} \cdot x_t^{t+1} = -p^m(x_t^{t+1,m} - m_t^{t+1}) + p^{t+1} \cdot \omega_t^{t+1} \qquad \textit{for} \quad t \geqslant 1. \end{aligned} \tag{3.4.1}$$

Proof. Equation (3.4.1) for $t = 0$ is derived from the System (2.2) and Proposition (3.1), using the facts that $u_0(\cdot)$ is strictly increasing and p^1 is positive. We next establish that Eqs. (3.4.1) hold for $t = 1$. Because $u_1(\cdot)$ is strictly increasing and p^2 is positive, the system (2.2) yields

$$p^1 x_1^1 + p^m(x_1^{1,m} - m_1^1) + p^2 \cdot x_1^2 = p^1 \cdot \omega_1^1 - p^m(x_1^{2,m} - m_1^2) + p^2 \cdot \omega_1^2$$

after applying Propositon (3.1). Adding the above equation to Eqs. (3.4.1) for $t = 0$ yields (3.4.1) for $t = 1$ after substituting the equalities

$$p^1 \cdot (x_0^1 + x_1^1) = p^1 \cdot (\omega_0^1 + \omega_1^1)$$

and

$$p^m(x_0^{1,m} + x_1^{1,m}) = p^m(m_0^1 + m_1^1),$$

which are consequences of equilibrium (Definition 2.3). The remainder of the proof is by induction on t. Assume that Eqs. (3.4.1) hold for $t = s - 1$. Then, by using the equilibrium conditions

$$\begin{aligned} & p^s \cdot (x_{s-1}^s + x_s^s) = p^s \cdot (\omega_{s-1}^s + \omega_s^s), \\ & p^m(x_{s-1}^{s,m} + x_s^{s,m}) = p^m(m_{s-1}^s + m_s^s), \end{aligned}$$

Eqs. (3.4.1) are established for $t = s > 1$. ∎

3.5. ***Remark.*** Notice that at equilibrium two separate "budget constraints" are satisfied for consumer $t \geqslant 1$, cf. (3.4.1). Thus, at equilibrium prices, consumer t can be thought of as purchasing and selling commodities solely on the current spot market, while providing for his future by savings held in the form of nonperishable money.

3.6. **Proposition.** *If x^m is a sequence of individual money additions with a constant present price of money $p^{t,m} = p^m > 0$ for $t = 1, 2,\ldots$, then*

$$\begin{aligned} &x_0^{1,m} = 0, && \\ &x_t^{t,m} = -x_t^{t+1,m} = m^t && \text{for} \quad t \geqslant 1, \qquad (3.6.1)\\ &x_t^{s,m} = 0 && \text{otherwise.} \end{aligned}$$

Proof. Proposition (3.1) allows us to set $p^{t,m} = p^m$. From Proposition (3.3), $x_0^{1,m} = 0$ and $x_t^{t,m} = -x_t^{t+1,m}$ for $t \geqslant 1$. From the equilibrium condition for clearing in the money market, we have that $x_t^{t,m} = m^t$ for $t \geqslant 1$. ▮

3.7. *Remark.* Proposition (3.6) reflects our simple demographic assumptions. In each period, there are only two consumers: An "old" consumer in his last period of life, and a "young" consumer with only one more period of life remaining. In equilibrium, the old consumer exchanges all of his outside money for goods of the young consumer. With this simple generational structure, there is in equilibrium no role for inside money; the money holdings of the young equal the stock of outside money, or government debt. Proposition (3.6) also provides a "dichotomy" between the money markets and the commodity markets. One should not however, make too much of the seeming separation of the economy into monetary and real parts. The demand for money is a correspondence depending on the sequence $p^{1,m}, p^{2,m},\ldots$. If for some s and t $(s \neq t)$, $p^{s,m} \neq p^{t,m}$, then the demand for money in some periods is arbitrarily large and the supply of money in some other periods is arbitrarily large. Furthermore, if $p^{s,m} \neq p^{t,m}$, then the demand for physical commodities is unbounded. If $p^{t,m} = p^m$ for $t = 1, 2,\ldots$, then individuals are indifferent as to how much money they should hold in their first periods of life. The constancy of the present price of money ensures that consumers are indifferent at the margin between a "dollar" of increased money holdings and a "dollar" of current consumption.

4. THE WALRASIAN EQUILIBRIUM EQUATIONS

We can derive from the system (2.2) individual demands for the physical commodities as functions of commodity prices $p \in \mathscr{S}$ and individual incomes $w \in W$. Thus, if f_t is the demand function for consumer t, then

$$\begin{aligned} &f_0\colon S \times \mathbb{R}_{++} \to \mathbb{R}_{++}^{l}, \\ &f_t\colon S \times \mathbb{R}_{++} \to \mathbb{R}_{++}^{2l} \qquad \text{for} \quad t \geqslant 1, \end{aligned} \qquad (4.1)$$

where

$$\begin{aligned} w_0 &= p \cdot \omega_0 + p^m m_0^1, \\ w_t &= p \cdot \omega_t + p^m(m_t^t + m_t^{t+1}) \qquad \text{for} \quad t \geqslant 1, \end{aligned} \tag{4.2}$$

when attention is restricted to the case where $p^{t,m} = p^m$. In this case a monetary policy affects individual wealth through (4.2) and thus affects individual demands, f_0 and f_t.

Notice from (4.1) and (4.2) that the effect of a monetary transfer $m \in M$ on individual demands for commodities depends only on the sums $m_0^1, m_1^1 + m_1^2, \ldots, m_t^t + m_t^{t+1}, \ldots$. This is an immediate consequence of perfect "capital" markets. A consumer is indifferent as to the timing of taxes and transfers as long as the total *present* value of his net tax bill is constant. We can thus rewrite (4.2) as

$$w_t = p \cdot \omega_t + p^m \mu_t \qquad \text{for} \quad t \geqslant 0, \tag{4.3}$$

where $\mu = (\mu_0, \mu_1, \ldots, \mu_t, \ldots) = (m_0^1, m_1^1 + m_1^2, \ldots, m_t^t + m_t^{t+1}, \ldots)$ is a *monetary policy*. Let $\mathscr{M}$ be the set of feasible government monetary policies μ. Since no restrictions are imposed on these policies, $\mathscr{M}$ is $\mathbb{R} \times \mathbb{R} \times \cdots = \mathbb{R}^\infty$, the case of full monetary control.

We next define Walrasian equilibrium and then establish the relationship between this equilibrium concept and the concept of monetary competitive equilibrium given in Definition (2.3).

4.4. DEFINITION. For given commodity endowments $\omega \in X$ and a given monetary policy $\mu \in \mathscr{M}$, a *Walrasian equilibrium* is a price system $q = (p, p^m) \in \mathscr{Q}$ which solves the following equations:

$$\sum_t f_t(p, w_t) = \sum_t \omega_t$$

and

$$w_t = p \cdot \omega_t + p^m \mu_t \qquad \text{for} \quad t \geqslant 0.$$

4.5. PROPOSITION. *Let $q = (p, p^m)$ be a Walrasian equilibrium for $\omega \in X$ and $\mu \in \mathscr{M}$. Then $p \in \mathscr{P}$ and the sequence $p^{1,m}, p^{2,m}, \ldots$, where $p^{t,m} = p^m$ for $t = 1, 2, \ldots$, is a monetary competitive equilibrium (Definition 2.1) for ω and $m = m_0^1, m_1^1, m_1^2, \ldots, m_t^t, m_t^{t+1}, \ldots$ if $\mu = (m_0^1, m_1^1 + m_1^2, \ldots, m_t^t + m_t^{t+1}, \ldots)$. Furthermore, the Walrasian allocation $f = (f_0, f_1, \ldots)$ is also the monetary competitive allocation, $f = x = (x_0, x_1, \ldots)$. If $p^m > 0$, competitive equilibrium money holdings are given by Proposition (3.9); if $p^m = 0$, any pattern of money holdings consistent with* *materials balance in the money market is also a competitive allocation of money.*

Proof. Follows directly (3.1), (3.3), (3.4), and (4.4). ▮

Proposition (4.5) allows us to loosely refer to "Walrasian equilibria" in what follows as "monetary competitive equilibria." Similarly, we refer to $\mu \in \mathscr{M}$ as a "monetary policy"; there should be no confusion with $m \in M$, also referred to as a "monetary policy."

4.6. DEFINITION. Let $\mathscr{Q}(\omega, \mu)$ denote the *set of Walrasian equilibria* (alternatively, the *set of monetary competitive equilibria*) for $\omega \in X$ and $\mu \in \mathscr{M}$. Thus $\mathscr{Q}(\omega, \mu) = \{q \in \mathscr{Q} \mid \text{Definition (4.4) is satisfied for } \omega \in X \text{ and } \mu \in \mathscr{M}\}$. The price system $q = (p, p^m)$ is said to define a *proper monetary competitive equilibrium* for $\omega \in X$ and $\mu \in \mathscr{M}$ if $q \in \mathscr{Q}(\omega, \mu)$ *and* $p^m \neq 0$.

4.7. PROPOSITION (The "Neutrality" of Money). *For each positive scalar* λ,

$$\mathscr{Q}(\omega, \lambda\mu) = \{p, (p^m/\lambda) \mid (p, p^m) \in \mathscr{Q}(\omega, \mu)\}.$$

That is, two economies which differ only in the units used for measuring money will have the same set *of equilibrium commodity prices and the same* set *of equilibrium* real *monetary transfers* $p^m\mu$.

Proof. From (4.3), $w_t = p \cdot \omega_t + p^m\mu_t$. Since the only effect of μ or p^m on f_t is through w_t, the result follows immediately. ▮

4.8. *Remark.* Proposition (4.7) is weaker than the usual "Quantity Theory of Money," which necessarily equates "doubling the nominal money supply" with "doubling the general price level." Here, if for *each period* the perfectly foreseen money transfer is doubled, then halving the present price of money is *consistent* with equilibrium, but is hardly necessary for equilibrium.

4.9. *Remark.* The constancy of the equilibrium price of money (Proposition (3.1)) depends on the assumption that the nominal rate of interest (the own rate of interest on money) is zero. To extend our analysis to the more general setting is quite simple. For example, if the nominal rate of interest is θ, the zero-profit condition becomes

$$\frac{p^{t+1,m}}{p^{t,m}} = \frac{1}{1+\theta},$$

so that a change in the nominal interest rate is exactly offset by an induced change in the inflation rate. In the more general setting, given ω the equilibrium allocation $f \in X$ is affected only by the sequence of real money transfers $(p^{1,m}m_0^1, \ldots, p^{t,m}m_t^t + p^{t+1,m}m_t^{t+1}, \ldots)$. As long as government policy does not affect this sequence, it does not affect the equilibrium allocation $f \in X$. In particular, if the government were to increase the nominal interest rate and the growth rate of nominal monetary transfers, all in the same proportion, then the set of equilibrium allocations would be unaffected. This simple property of the model is known as the *superneutrality of money*.

5. Bonafide Monetary Policies

From Definition (4.4), a Walrasian equilibrium $q = (p, p^m)$ is determined (not necessarily uniquely) by the demand functions, the endowments, and the government monetary policy. Fixing the demand functions f, if the government selects monetary policy μ when the endowments of the economy are ω, then we say that μ is an *ω-monetary policy*.

The purpose of monetary policy is to affect the equilibrium allocation $x = f$. If, for some nontrivial monetary policy μ, all Walrasian equilibria associated with the endowments ω have the property that $p^m = 0$, then the monetary policy μ cannot possibly have any affect on the Walrasian allocation, $x = f$. In this case, the government could have no "*good-faith*" expectation that its monetary policy would be effective. The monetary policy μ is then not *bonafide* for the economy with endowments ω.

5.1. Definition. The ω-monetary policy $\mu \in \mathscr{M}$ is said to be *ω-bonafide* if there is a *proper* monetary equilibrium $q = (p, p^m) \in \mathscr{P} \times \mathbb{R}_{++}$ associated with (ω, μ), i.e., there is a $q = (p, p^m) \in \mathscr{Q}(\omega, \mu)$ such that $p^m \neq 0$. The monetary policy $\mu \in \mathscr{M}$ is said to be *bonafide* if it is ω-bonafide for some $\omega \in X$.

Clearly, the existence of ω-bonafide and bonafide monetary policies is critical for the integration of money within the general-equilibrium model of overlapping generations.[1] We investigate in some detail the properties of these policies.

5.2. Proposition. *The set of ω-bonafide monetary policies and the set of bonafide monetary policies are cones.*

---

[1] The finite-horizon model is unsatisfactory for monetary analysis because only very special monetary policies permit a nonzero price of money. In the finite-horizon, n-consumer model, only policies in which the algebraic sum of taxes is zero, $\sum_1^n \mu_i = 0$, are bonafide. See, e.g., Balasko and Shell [7].

Proof. If μ is ω-bonafide, then by Definitions (4.4) and (5.1), there is $q = (p, p^m) \in \mathscr{Q}$ with $p^m \neq 0$ such that $q \in \mathscr{Q}(\omega, \mu)$. Then by Proposition (4.7), $(p, p^m/\lambda) \in \mathscr{Q}(\omega, \lambda\mu)$, where $\lambda > 0$. ∎

Thus, the neutrality of money guarantees that the set of ω-bonafide monetary policies and hence the set of bonafide monetary policies are cones. It suffices, therefore, to analyze convenient cross sections of these sets; such normalizations are provided in the following definition.

5.3. Definition. The ω-bonafide (resp., bonafide) monetary policy $\mu \in \mathscr{M}$ is said to be *normalized* if $(p, 1)$ is a (proper) monetary competitive equilibrium associated with (ω, μ) (resp., with (ω, μ) for some ω). We denote by $\mathscr{M}_B(\omega) \subset \mathscr{M}$ (resp., $\mathscr{M}_B \subset \mathscr{M}$) the *set of normalized ω-bonafide* (resp., *normalized bonafide*) *monetary policies.* The set of ω-bonafide monetary policies is the positive cone in $\mathscr{M}$ generated by $\mathscr{M}_B(\omega)$ (resp., $\mathscr{M}_B$).

In what follows, we characterize the set of normalized ω-bonafide monetary policies in terms of the price sequences $q = (p, 1) \in \mathscr{Q}$ and income sequences $w = w_0, w_1, \ldots$ consistent with the existence of (proper) monetary competitive equilibria. To do this, we extend to the overlapping-generations context, the price-income equilibrium concept employed in standard equilibrium theory (see, e.g., Balasko [3]).

The resource (or aggregate endowment) sequence $r = \sum_t \omega_t$ is taken as fixed. We are interested in the set of prices and incomes (p, w) that solve $\sum_t f_t(p, w_t) = r$. Let μ be an ω-bonafide normalized monetary policy where $(p, 1)$ is the associated monetary competitive equilibrium, so that $w_t = p \cdot \omega_t + \mu_t$ for $t \geqslant 0$. Given the normalization $p^m = 1$, we also ask whether or not μ is consistent with a given price-income equilibrium (p, w). These ideas are formalized in the following definition.

5.4. Definition. The sequence (p, w) is a *price-income equilibrium* associated with the total resources $r = (r^1, \ldots, r^t, \ldots)$ if and only if $\sum_t f_t(p, w_t) = r$. The ω-monetary (resp., monetary) policy μ is said to be *consistent with the price-income equilibrium* (p, w) if and only if $w_t = p \cdot \omega_t + \mu_t$ for $t \geqslant 0$ (resp., there is an ω such that $\sum_t \omega_t = r$ and $w_t = p \cdot \omega_t + \mu_t$ for $t \geqslant 0$).

Note that if μ is consistent with (p, w), then clearly $(p, 1)$ is a (proper) monetary equilibrium associated with (ω, μ) for some ω; in other words, μ is appropriately normalized.

The next proposition provides a characterization of the monetary policies consistent with a *given* price-income equilibrium.

5.5. **Proposition.** *The (normalized) monetary policy μ is consistent with the price-income equilibrium (p, w) if and only if the following system of inequalities is satisfied*

$$-p^{t+1}\cdot f_{t+1}^{t+1}(p, w_{t+1}) < \sum_{l=0}^{t} \mu_l < p^{t+1}\cdot f_t^{t+1}(p, w_t)$$

$$\text{for} \quad t \geqslant 0. \qquad (5.5.1)$$

Proof. *Step* 1. Consider the equilibrium equations

$$f_0^1(p, w_0) + f_1^1(p, w_1) = r^1,$$

$$\cdots \qquad \cdots \qquad = \cdots$$

$$f_t^{t+1}(p, w_t) + f_{t+1}^{t+1}(p, w_{t+1}) = r^{t+1}.$$

Note that we consider only the first $(t+1)$ equilibrium equations. Multiply equation i by p^i, and add up. Applying Walras' law to each individual demand function, we obtain

$$w_0 + w_1 + \cdots + w_t + p^{t+1}\cdot f_{t+1}^{t+1} = p^1\cdot r^1 + \cdots + p^{t+1}\cdot r^{t+1}.$$

Therefore, we have

$$\sum_{l=0}^{t} w_l - \sum_{l=1}^{t+1} p^l\cdot r^l = -p^{t+1}\cdot f_{t+1}^{t+1}$$

and

$$\sum_{l=0}^{t} w_l - \sum_{l=1}^{t} p^l\cdot r^l = p^{t+1}\cdot (r^{t+1} - f_{t+1}^{t+1}) = p^{t+1}\cdot f_t^{t+1}.$$

We can replace the left-hand side and the right-hand side of (5.5.1) by the above expressions, yielding

$$\sum_{l=0}^{t} w_l - \sum_{l=1}^{t+1} p^l\cdot r^l < \sum_{l=0}^{t} \mu_l < \sum_{l=0}^{t} w_l - \sum_{l=1}^{t} p^l\cdot r^l$$

$$\text{for} \quad t \geqslant 0. \qquad (5.5.2)$$

Step 2. We now determine explicitly the sequence of endowments $\omega = (\omega_0, \ldots, \omega_t, \ldots)$ such that $(p, 1)$ is a monetary equilibrium associated with (ω, μ).

Start with $\omega_0 = \omega_0^1$. Clearly, we must have $p^1\cdot\omega_0^1 = w_0 - \mu_0$. The inequality (5.5.2) reduces, in the case $t = 0$, to $w_0 - p^1\cdot r^1 < \mu_0 < w_0$. We want an endowment ω_0^1 such that $0 < \omega_0^1 < r^1$. Such a solution exists if $0 < p^1\cdot\omega_0^1 < p^1\cdot r^1$, i.e., $0 < w_0 - \mu_0 < p^1\cdot r^1$ or, equivalently,

$w_0 - p^1 \cdot r^1 < \mu_0 < w_0$. Therefore, (5.5.2) for $t = 0$ provides us with ω_0^1 such that

$$p^1 \cdot \omega_0^1 = w_0 - \mu_0$$
$$0 < \omega_0^1 < r^1.$$

Let $\omega_1^1 = r^1 - \omega_0^1$. We clearly have $0 < \omega_1^1 < r^1$. Let us now determine ω_1^2 such that

$$p^1 \cdot \omega_1^1 + p^2 \cdot \omega_1^2 = w_1 - \mu_1,$$
$$0 < \omega_1^2 < r^2.$$

We have

$$\begin{aligned} p^2 \cdot \omega_1^2 &= w_1 - \mu_1 - p^1 \cdot \omega_1^1 \\ &= w_1 - \mu_1 - p^1 \cdot (r^1 - \omega_0^1) \\ &= w_1 - \mu_1 - p^1 \cdot r^1 + w_0 - \mu_0 \\ &= w_0 + w_1 - p^1 \cdot r^1 - \mu_0 - \mu_1. \end{aligned}$$

There is an ω_1^2 satisfying

$$p^2 \cdot \omega_1^2 = w_0 + w_1 - p^1 \cdot r^1 - \mu_0 - \mu_1,$$
$$0 < \omega_1^2 < r^2,$$

if and only if $0 < p^2 \cdot \omega_1^2 < p^2 \cdot r^2$, i.e.,

$$0 < w_0 + w_1 - p^1 \cdot r^1 - \mu_0 - \mu_1 < p^2 \cdot r^2,$$

or, equivalently,

$$w_0 + w_1 - p^1 \cdot r^1 - p^2 \cdot r^2 < \mu_0 + \mu_1 < w_0 + w_1 - p^1 \cdot r^1,$$

which corresponds to (5.5.2) for $t = 1$.

The rest of the proof proceeds by induction. Therefore, we can assume that ω_t^t is determined if and only if the inequality (5.5.2) for $(t-1)$ holds. We also have

$$p_t \cdot \omega_t^t = -\sum_{l=0}^{t-1} w_l + \sum_{l=1}^{t} p^l \cdot r^l + \sum_{l=0}^{t-1} \mu_l.$$

Therefore, the problem is to find ω_t^{t+1} such that

$$w_t = p^t \cdot \omega_t^t + p^{t+1} \cdot \omega_t^{t+1} + \mu_t,$$
$$0 < \omega_t^{t+1} < r^{t+1}.$$

The first equality becomes

$$p^{t+1} \cdot \omega_t^{t+1} = \sum_{l=0}^{t} w_l - \sum_{l=1}^{t} p^l \cdot r^l - \sum_{l=0}^{t} \mu_l .$$

Therefore, there is an ω_t^{t+1} satisfying the above constraint if and only if

$$0 < p^{t+1} \cdot \omega_t^{t+1} < p^{t+1} \cdot r^{t+1},$$

which, as before, is easily transformed into

$$0 < \sum_{l=0}^{t} w_l - \sum_{l=1}^{t} p^l \cdot r^l - \sum_{l=0}^{t} \mu_l < p^{t+1} \cdot r^{t+1},$$

which corresponds to (5.5.2) for $t > 1$. ∎

The "only if" part of the proof of Proposition (5.5) is very easy. An alternative "only if" proof follows directly from Propositions (3.4) and (3.6). Given p, w, and μ, there is a positive ω_t^{t+1} for $t \geqslant 0$ only if

$$p^{t+1} \cdot f_t^{t+1}(p, w_t) > m^t + m_t^{t+1} = \sum_{l=0}^{t} \mu_l .$$

Also from Propositions (3.4) and (3.6), there is a positive ω_{t+1}^{t+1} for $t \geqslant 0$ consistent with p, w, and μ only if

$$\sum_{l=0}^{t} \mu_l > -p^{t+1} \cdot f_{t+1}^{t+1}(p, w_{t+1}).$$

Combining the inequalities yields the system (5.5.1).

5.6. *Remark.* The trivial monetary policy $\mu = 0$ obviously satisfies the inequality system (5.5.1). This is not surprising since the price-income equilibrium (p, w) can be associated with $(\omega, 0)$, where $\omega_t = f_t(p, w_t)$ for $t \geqslant 0$.

5.7. PROPOSITION. *Let* $(p, 1)$ *be a monetary competitive equilibrium associated with* (ω, μ), *i.e.*, $(p, 1) \in \mathscr{Q}(\omega, \mu)$. *If the total resources* $r = \sum_t \omega_t$ *are bounded from above (i.e., there is* $0 < K < +\infty$ *such that* $\|r^t\| < K$ *for* $t \geqslant 1$), *then*

$$\|p^t\| > \frac{1}{K} \left\| \sum_{l=0}^{t-1} \mu_l \right\| \qquad \text{for} \quad t \geqslant 1,$$

where for convenience we use the Euclidean norm.

Proof. From (5.5.1), we deduce that

$$\left\| \sum_{l=0}^{t-1} \mu_l \right\| < \max(\|p^t \cdot f^t_{t-1}\|, \|p^t \cdot f^t\|) \leqslant \|p^t \cdot r^t\| \leqslant \|p^t\| \, \|r^t\|,$$

so that

$$\|p^t\| > \frac{1}{\|r^t\|} \left\| \sum_{l=0}^{t-1} \mu_l \right\| > \frac{1}{K} \left\| \sum_{l=0}^{t-1} \mu_l \right\|. \quad \blacksquare$$

Proposition (5.7) is important since it relates the dynamic behavior of the "money supply," $\{\sum_{l=0}^{t} \mu_l\}$, with properties of the sequence of commodity prices, $\{p^t\}$, which properties are known to be critical in determining whether or not the competitive allocation $\{f_t(p, w_t)\}$ is Pareto-optimal. (See Balasko and Shell [5, Sect. 5].)

First, we apply Proposition (5.7) to an important class of equilibria satisfying a long-run "positive interest-rate" condition: $\|p^t\| \to 0$ as $t \to \infty$. We know that such equilibria are Pareto-optimal if the allocation sequence $\{x_t\}$ is bounded from above; cf. [5, Proposition (5.3)].

5.8. PROPOSITION. *Assume*

> *Property* B: *The allocation sequence* $\{x_t\}$ *is bounded from above, so that there is* $0 < K < +\infty$ *such that* $\|r^t\| < K$ *for* $t \geqslant 1$.

Let $\mu = (\mu_0, \mu_1, ..., \mu_t, ...) \in \mathscr{M}$ *be a bonafide monetary policy admitting the proper monetary competitive equilibrium* (p, p^m) *(so that* $p^m \neq 0$*). If* $\|p^t\|$ *tends to zero as* t *tends to infinity, then* μ *necessarily satisfies the equality* $\sum_{t=0}^{\infty} \mu_t = 0$.

Proof. If $\mu \in \mathscr{M}_B(\omega)$ admits the monetary equilibrium $(p, 1)$, then from Proposition (5.7), we have that

$$\left\| \sum_{l=0}^{t-1} \mu_l \right\| < K \|p^t\| \qquad \text{for} \quad t \geqslant 1.$$

Since $\|p^t\| \to 0$ as $t \to \infty$ the right-hand side of this inequality approaches zero as t approaches infinity. Hence, the series $\{\sum_0^{t-1} \mu_l\}$ converges and $\sum_0^{\infty} \mu_t = 0$. Proposition (5.2) allows this result to be extended to (non-normalized) bonafide monetary policies. $\blacksquare$

Proposition (5.8) has a simple economic interpretation. If the present prices of commodities are falling to zero, the *current* price of money in terms of commodities is becoming arbitrarily large. Unless the endowment of commodities is growing, this situation is inconsistent with monetary equilibrium: ultimately the demand for commodities must exceed their supply.

From Proposition (5.8), we conclude that a "positive-interest-rate" economy with a nonpositive long-run real growth rate allows only a limited set of bonafide monetary policies, i.e., those satisfying the restriction that the money supply tends to zero as t becomes large. Such economies are like finite-horizon economies in that for a monetary policy to be bonafide, the algebraic sum of transfers must be zero; cf. [7]. Proposition (5.8) generalizes to the case of real economic growth (i.e., growth in resources, r^t): If the long-run interest rates exceed the long-run real growth rates, then the long-run money supply must be zero for any ω-bonafide monetary policy. On the other hand, if the rate of monetary growth is greater than the real rate of growth of the economy, then Pareto efficiency is ruled out, as the next proposition shows for an economy in which consumers satisfy some mild uniformity properties.

5.9. **Proposition.** *Consider the monetary policy $\mu \in \mathscr{M}$ satisfying the growth condition*

$$\sum_{s=0}^{t} \mu_s \geqslant \mu_0 (1+\gamma)^t \qquad \textit{for} \quad t \geqslant 0,$$

where γ is a positive scalar. Let $q = (p, p^m)$ be a proper monetary competitive equilibrium associated with μ and some $\omega \in X$, i.e., $(p, p^m) \in \mathscr{Q}(\omega, \mu)$ and $p^m \neq 0$ and let $x = \{x_t\} = \{f_t(p, p \cdot \omega_t + p^m \mu_t)\} \in X$ be the corresponding equilibrium allocation. Assume, furthermore, that consumers satisfy uniformity properties (a–e) *of Proposition* (5.6) *in Balasko and Shell* [5].[2] *Then, the allocation x is not Pareto-optimal.*

Proof. Since we include in our uniformity conditions the boundedness of the allocation sequence,we know from Proposition (5.7) that there is a positive scalar K such that

$$\|p^t\| > \frac{1}{K} \mu_0 (1+\gamma)^{t-1} \qquad \text{for} \quad t \geqslant 1,$$

in the normalized case, where $p^m = 1$. It follows from Balasko and Shell [4], Proposition (5.6), that the resulting normalized monetary competitive equilibrium allocation, $x = f(p, w)$ is not Pareto-optimal. From

---

[2] A close reading of the proof of Proposition (5.6) in [5] shows that not all of the uniformity properties (a–e) are needed for Proposition (5.9). In particular, while the upper bound on the Gaussian curvature of indifference surfaces is crucial, we do not require here that the curvature be bounded above zero.

Proposition (5.2), we immediately deduce that the corollary applies to all proper monetary competitive equilibria, i.e., for any positive p^m. ▮

Our result can be taken as a generalization of a single-commodity steady-state result in Wallace [19] (cf. especially his Proposition 6), a provocative paper full of insights and results. Our Propositions (5.8) and (5.9) along with Proposition (5.6) in [5] strongly suggest that attention be focused on monetary policies for which the asymptotic growth rate of the money supply is equal to "the" asymptotic growth rate of the real economy. The equation of these rates of growth is consistent with some of the prescriptions of the Chicago School.[3] Nothing in our analysis, however, supports a restriction to a fixed short-run growth rate for money.

6. THE SET OF ω-BONAFIDE MONETARY POLICIES

Proposition (5.5), the central result of the preceding section, provides a characterization of those bonafide monetary policies which are consistent with a given price-income equilibrium (p, w). This characterization aids us in recognizing whether or not a monetary policy $\mu \in \mathscr{M}$ is bonafide and consistent with (p, w) for some suitably assigned sequence of endowments of commodities $\omega \in X$.

These results are not, however, sufficient for the analysis of money as a policy instrument. For example, the government might consider choosing a monetary policy that would lead to allocations which are Pareto-superior to those attainable without money. Or, more generally, the government might attempt to promote through monetary policy an allocation which maximizes some interpersonal welfare or distributional criterion. In pursuing these goals, the government may well have no power to alter the endowments of physical commodities, $\omega \in X$, but is free to employ monetary policy, $\mu \in \mathscr{M}$, as a control variable.

In this section, we analyze properties of the set of ω-bonafide monetary policies. The first problem we consider is whether each Pareto-optimal (PO) allocation, $x \in X$, can be achieved as a monetary competitive allocation associated with the given endowments, $\omega \in X$ $(\sum_t x_t = \sum_t \omega_t)$, and some $\mu \in \mathscr{M}_B(\omega)$, the set of normalized bonafide monetary policies.

6.1. PROPOSITION. *There is one and only one monetary policy $\mu \in \mathscr{M}_B(\omega) \subset \mathscr{M}$ such that a given Pareto-optimal allocation $x = \{x_t\}$ is a*

---

[3] A precise citation to the oral tradition of the Chicago School is often difficult or impossible. For some clues, see Friedman [12, especially pp. 133–156]. The emphasis in [12] is on fixed rules. The desirability of equating real and monetary growth rates is, at best, implicit in [12].

competitive allocation associated with (ω, μ) *for some monetary competitive equilibrium* $(p, 1)$, *i.e., where* $x_t = f_t(p, w_t)$ *and* $w_t = p \cdot \omega_t + \mu_t$ *for* $t \geqslant 0$.

Proof. Since x is Pareto optimal (PO) and thus weakly Pareto optimal (WPO), there is a unique price sequence $p \in \mathscr{P}$ supporting x, i.e., $x = f_t(p, p \cdot x_t)$ for $t \geqslant 0$. (See [5, Lemma (4.3)].) Let $w_t = p \cdot x_t$ for $t \geqslant 0$. We have that $\sum_t f_t(p, w_t) = r$, and hence (p, w) is a price-income equilibrium (Definition (5.4)). To obtain x as the competitive allocation associated with (ω, μ) for the (proper) monetary competitive equilibrium $(p, 1)$, we merely require that $w_t = p \cdot \omega_t + \mu_t$, or $\mu_t = w_t - p \cdot \omega_t$, for $t \geqslant 0$. Given $x \in X$, $p \in \mathscr{P}$ is unique, hence $w \in W$ is unique, so that $\mu \in \mathscr{M}$ (as constructed above) exists and is unique. ▮

A result close to our Proposition (6.1) was established by Okuno and Zilcha in [15, Theorem 2].

To achieve the PO allocation x through monetary policy alone is always possible, but it is not assured. The Pareto-optimizing μ is unique in $\mathscr{M}_B(\omega)$, the set of *normalized* ω-bonafide monetary policies. The Pareto-optimizing ω-bonafide monetary policy, however, is in general of the form (μ/p^m) where the monetary competitive equilibrium is $(p, p^m) \in \mathscr{P} \times \mathbb{R}_{++}$. Thus the scale of the Pareto-optimizing ω-bonafide monetary policy is indeterminate. There is no assurance that the competitive economy will choose the present price of money (given μ and ω) to achieve the intended PO allocation. Proposition (6.1) then has the interpretation that to ensure a PO allocation x, the government must in general control both $\mu \in \mathscr{M}$ and $p^m \in \mathbb{R}_+$.

Let $\mathscr{B}$ (for "budget") denote the set of price-income equilibria, i.e.,

$$\mathscr{B} = \left\{ (p, w) \in \mathscr{P} \times W \mid w_t > 0 \text{ for } t \geqslant 0 \text{ and } \sum_t f_t(p, w_t) = r \right\},$$

where r, the sequence of total resources, is given and held constant. Let Φ_ω: $\mathscr{B} \to \mathscr{M}$ denote the mapping given by

$$(p, w) \mapsto (\mu_0 = w_0 - p \cdot \omega_0, \ldots, \mu_t = w_t - p \cdot \omega_t, \ldots).$$

6.2. PROPOSITION. $\mathscr{M}_B(\omega) = \Phi_\omega(\mathscr{B}) = \operatorname{Im} \Phi_\omega$.

Proof. Obvious. ▮

Proposition (6.2) is a reflection of the simple structure of equilibrium in the money markets (cf. Sections 3 and 4). Proposition (6.2) simplifies the analysis of the set of ω-bonafide monetary policies, since variations of μ in $\mathscr{M}_B(\omega)$ can be related to variations of (p, w) in $\mathscr{B}$.

6.3. PROPOSITION. *The monetary policy sequence* $0 \in \mathscr{M}$ *belongs to the set* $\mathscr{M}_B(\omega)$.

Proof. Let $p \in \mathscr{P}$ be an equilibrium price sequence associated with the *nonmonetary* economy with endowments $\omega \in X$ (cf. [5, Definition (2.3)]). Let $x = \{x_t\}$ be the corresponding competitive allocation. By definition, $w_t = p \cdot x_t = p \cdot \omega_t$ for $t \geqslant 0$ and hence $\Phi_\omega(p, w) = 0$ if $w = \{w_t\} = \{p \cdot \omega_t\}$. ∎

In general, $\mathscr{M}_B(\omega)$ is a proper subset of $\mathscr{M}$. There can be monetary policies $\mu \in \mathscr{M}$ which are not ω-bonafide; see, for example, the highly regular "classical-case" examples in Gale [13]. Regular examples can also be constructed in which the sequence 0 does not necessarily belong to the interior of $\mathscr{M}_B(\omega)$. (This last contrasts with our analysis of the regular *finite* economy, where in fact 0 belongs to the interior of the set of ω-bonafide monetary (or, alternatively, tax) policies; see Balasko and Shell [7].)

An interesting question which arises in the analysis of economic policy is whether the control instruments can be altered "continuously" within some set. We investigate here whether it is possible to move in a continuous way from one ω-bonafide monetary policy to another. The continuity properties of $\mathscr{M}_B(\omega)$ turn out to enable us to deduce properties of the set of monetary competitive equilibria; see Section 7. We also investigate the continuity of the set of monetary policies consistent with Pareto-optimality.

The natural concept of continuity in this context turns out to be that of connectedness.

6.4. Proposition. *The set $\mathscr{M}_B(\omega)$ is connected.*

Proof. The mapping $\Phi_\omega \colon \mathscr{P} \times W \to \mathscr{M}$ is continuous. The set of price-income equilibria $\mathscr{B}$ being connected (see Corollary (A.1.3) in the Appendix), its image $\Phi_\omega(\mathscr{B})$ is, therefore, connected (cf. Bourbaki [8, Chap. I, Sect. 11.3, Proposition 4]). The result follows after applying Proposition (6.2). ∎

The fact that PO allocations are potentially implementable through some bonafide monetary policies leads us to introduce the concept of a *potent* normalized monetary policy.

6.5. Definition. A monetary policy $\mu \in \mathscr{M}_B(\omega)$ is said to be *potent* if there exists a monetary competitive equilibrium $(p, 1)$ associated with (ω, μ) such that the allocation $x = \{x_t\} = \{f_t(p, p \cdot \omega_t + \mu_t)\}$ is Pareto-optimal.

Let $\mathscr{M}_P(\omega)$ denote the set of potent ω-monetary policies. Clearly, it follows that $\mathscr{M}_P(\omega) \subset \mathscr{M}_B(\omega) \subset \mathscr{M}$. Note that the sequence 0 does not necessarily belong to $\mathscr{M}_P(\omega)$; see, for example, the "Samuelson-case" example in Gale [13].

Let $\mathscr{B}_P$ denote the subset of the set of price-income equilibria which correspond to Pareto-optimal allocations, i.e., for which $\{f_t(p, w_t)\}$ is Pareto-optimal. It follows that $\mathscr{M}_P(\omega) = \Phi_\omega(\mathscr{B}_P)$.

6.6. **Proposition.** *The set $\mathscr{M}_p(\omega)$ is connected.*

Proof. From Proposition (A.2.2) in the Appendix, it follows that $\mathscr{B}_p$ is arc-connected and hence connected. Therefore, $\Phi_\omega(\mathscr{B}_p) = \mathscr{M}_p(\omega)$ is connected. ▮

7. The Set of Equilibrium Money Prices

Let $\mathscr{Q}(\omega, \mu)$ be the set of monetary competitive equilibria $q = (p, p^m)$ for given ω and μ. (Note that μ is neither necessarily ω-bonafide nor necessarily normalized.) The *set of equilibrium (present) money prices* associated with the monetary competitive equilibrium is denoted by

$$\mathscr{P}^m(\omega, \mu) = \{ p^m \mid (p, p^m) \in \mathscr{Q}(\omega, \mu) \} \subset \mathbb{R}_+ .$$

In this section, we establish some basic properties of the set $\mathscr{P}^m(\omega, \mu)$.

7.1. **Proposition.** *The set $\mathscr{P}^m(\omega, \mu)$ is not empty.*

Proof. Take $p^m = 0$. Then $q = (p, 0)$ is a monetary competitive equilibrium if and only if p is an equilibrium of the nonmonetary economy defined by the endowment sequence ω. From [5], Proposition (3.10), we know that such an equilibrium sequence p exists. Then $(p, 0) \in \mathscr{Q}(\omega, \mu)$ and $0 \in \mathscr{P}^m(\omega, \mu)$. ▮

Proposition (7.1), although important, merely restates in the monetary framework the existence of equilibrium in the nonmonetary economy. It reminds us that fiat money's value is crucially dependent on the "faith" of consumers. No matter what monetary policy is pursued—bonafide or not—if consumers mistrust government debt it may well bear a zero value in equilibrium. Proposition (7.1) should be sharply contrasted with other monetary theories which claim a unique general price level which is positive and finite. It is not enough to assert that the community would benefit from a positive p^m if consumers refuse to accept that price as an equilibrium value.

We go on to investigate further the set $\mathscr{P}^m(\omega, \mu)$, now focusing on nontrivial equilibria.

7.2. **Proposition.** *$\mathscr{P}^m(\omega, \mu) \neq \{0\}$ if and only if μ is an ω-bonafide monetary policy.*

Proof. Let $p^m \in \mathscr{P}^m(\omega, \mu)$ be such that $p^m \neq 0$. Clearly, the sequence $p^m \mu \in \mathscr{M}$ is an ω-bonafide normalized monetary policy. ▮

Fixing the monetary policy $\mu \in \mathcal{M}$, define

$$L(\mu) = \{\lambda\mu \mid \lambda \geqslant 0\} \subset \mathcal{M},$$

the ray generated by μ.

7.3. **Proposition.** *Consider a nontrivial monetary policy $\mu \neq 0$ in $\mathcal{M}$. The set of equilibrium money prices $\mathcal{P}^m(\omega, \mu)$ is related to the set $\mathcal{M}_B(\omega) \cap L(\omega)$ by a one-to-one mapping.*

Proof. Map p^m from $\mathcal{P}^m(\omega, \mu)$ to $p^m\mu \in \mathcal{M}_B(\omega)$. This mapping is clearly a bijection, naturally arising because of the Neutrality of Money (Proposition (4.7)). ∎

The sets $\mathcal{M}_B(\omega)$ and $L(\mu)$ are depicted in Fig. 7.1. The plane of the page is intended to portray the infinite dimensional space $\mathcal{M}$ and the intersection of the axes is the point $\mu = 0$. The shaded set depicts $\mathcal{M}_B(\omega)$. $L(\mu)$ is the ray through the point μ. The set $\mathcal{M}_B(\omega) \cap L(\mu)$ is thus depicted by the two heavy line segments.

If the set $\mathcal{M}_B(\omega)$ were convex or even star-shaped about the origin, then the $\mathcal{M}_B(\omega) \cap L(\mu)$ would be an interval and thus $\mathcal{P}^m(\omega, \mu)$ would be an interval. In the general connected case, however, one can expect the set $\mathcal{M}_B(\omega) \cap L(\mu)$ to exhibit a rather complicated structure.

The set $\mathcal{P}^m(\omega, \mu)$ can be reduced to $\{0\}$ or to a proper interval; see [6, Proposition (3.2)]. We do not have a regular example where $\mathcal{P}^m(\omega, \mu)$ is a disconnected subset of $\mathbb{R}_+$, but it seems likely that one can construct such an example. Nonetheless, one cannot in general expect to have $\mathcal{P}^m(\omega, \mu) = \{0\} \cup \{p^{m*}\}$, where p^{m*} is the unique positive price of money. This has an important theoretical consequence: For a bonafide monetary policy, there is typically a (vast) indeterminacy of the general price level and a consequent indeterminacy of the resulting competitive allocation. This prevents the elaboration of a utility theory of money along the lines of Patinkin [15, Chaps V–VIII], where it is implied that the determination of p^m is unique at some positive level. In particular, Balasko and Shell [6] shows that for the log-linear case $\mathcal{P}^m(\omega, \mu)$ is an interval, so that for monetary models one cannot in general expect the equilibria to be isolated.

The set of equilibrium money prices may not be bounded from above. In particular, if the monetary policy is trivial, $\mu = 0$, then $\mathcal{P}^m(\omega, \mu) = \mathbb{R}_+$ for each $\omega \in X$. If the present price of money is arbitrarily large, then the current "price level" must be arbitrarily close to zero. This strange set of affairs is ruled out for usual monetary policies (those in which the first government action is to create a positive money stock). This is shown in the next proposition.

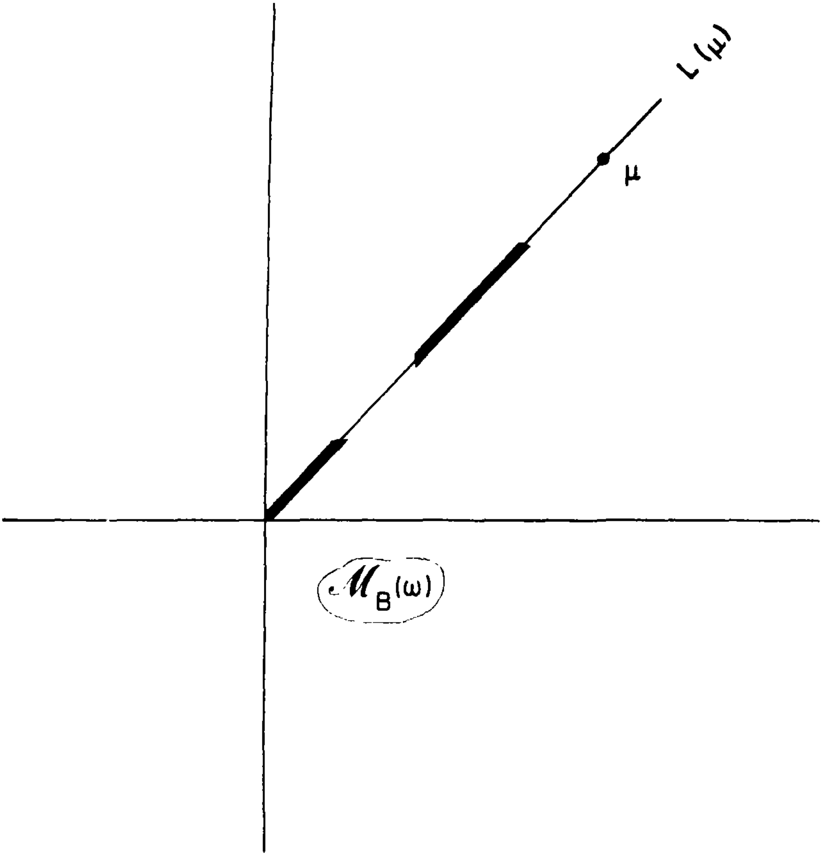


FIGURE 7.1

7.4. PROPOSITION. *Let* $\mu = (\mu_0, \mu_1, \ldots, \mu_t, \ldots) \in \mathcal{M}$. *Assume*:

(a) *For some* t, $\mu_t \neq 0$, *and*

(b) $\mu_{t_0} > 0$ *where* t_0 *is the first* t *with* $\mu_t \neq 0$,

Then, the set of equilibrium money prices $\mathcal{P}^m(\omega, \mu)$ *for the economy* (ω, μ) *is bounded if* ω *is bounded.*

The second hypothesis of Proposition (7.4) rules out pathological cases where the first nonzero monetary action of the government leads to the creation of a negative aggregate money supply.

Proof. The equilibrium allocation $x = (x_0, x_1, \ldots, x_{t_0}, x_{t_0+1})$ must satisfy the inequalities

$$\begin{aligned} u_0(x_0) &\geqslant u_0(\omega_0) \quad \text{and} \quad x_0 \leqslant r^1 && \text{for } t = 0, \\ u_t(x_t) &\geqslant u_t(\omega_t) \quad \text{and} \quad x_t \leqslant (r^t, r^{t+1}) && \text{for } t = 1, \ldots, t_0. \end{aligned} \tag{7.4.1}$$

Let (p, p^m) by a monetary competitive equilibrium. Because of the regularity of utility functions, we have from Inequalities (7.4.1) that prices $(p^1, \ldots, p^{t_0}, p^{t_0+1})$ are bounded above zero and bounded from above; cf. [5, Lemma (3.4)]. In particular there is $\beta^{t_0+1} \in \mathbb{R}^l_{++}$ such that $p^{t_0+1} \leqslant \beta^{t_0+1} < +\infty$. But, from Proposition (3.4), we have

$$p^{t_0+1} \cdot x^{t_0+1}_{t_0+1} + p^m(x^{t_0+1,m}_{t_0+1} - m^{t_0+1}_{t_0+1}) = p^{t_0+1} \cdot \omega^{t_0+1}_{t_0+1}.$$

From Proposition (3.6), we have $x^{t_0+1,m}_{t_0+1} = m^{t_0+1}$, so $x^{t_0+1,m}_{t_0+1} - m^{t_0+1}_{t_0+1} = \mu_{t_0}$. Hence we have

$$p^m \mu_{t_0} < p^{t_0+1} \cdot \omega^{t_0+1}_{t_0+1} \leqslant \beta^{t_0+1} \cdot \omega^{t_0+1}_{t_0+1} < +\infty,$$

since $\omega^{t_0+1}_{t_0+1}$ is bounded. The proposition then follows because μ_{t_0} is positive. ∎

APPENDIX

The purpose of this appendix is to establish connectedness (cf. Bourbaki [7, Chap. I, Sect. 11, Definition 1]) of the set of Pareto optima (PO) and the set of weak Pareto optima (WPO) in the pure-exchange, overlapping-generations model of Balasko and Shell [5]. These connectedness properties are interesting for their own sake and play a crucial role in the study of bonafide monetary policies and of potent monetary policies (cf. Section 6 of this paper, especially Propositions (6.4) and (6.6)).

A.1. *The Set of Weak Pareto Optima*

Let $\mathfrak{P}$ denote the set of weak Pareto optima (cf. Balasko and Shell [5, Definition 2.5]). The relationship between WPO allocations and competitive allocations has been clarified in [5, Proposition (4.4)]. We reformulate this relationship in the following form, where X, the space of feasible allocations, and $\mathscr{P} \times W$ are endowed with the product topology.

A.1.1. PROPOSITION. *The set of price-income equilibria* $\mathscr{B} = \{(p, w_0, \ldots, w_t, \ldots) \in \mathscr{P} \times W \mid \sum_t f_t(p, w_t) = r\}$ *is homeomorphic to the set of weak Pareto optima* $\mathfrak{P}$.

A.1.2. **Proposition.** *The set of weak Pareto optimal* $\mathfrak{P}$ *is connected.*

A.1.3. **Corollary.** *The set of price-income equilibria* $\mathscr{S}$ *is connected.*

Proof of (A.1.3). Obvious from (A.1.1) and (A.1.2). ∎

Proof of (A.1.2). Several steps are required.

Step 1. Define the set $\mathfrak{P}_n(t)$ (where n and t are natural numbers) as the set of allocations $x = (x_0, x_1, \ldots, x_t, \ldots) \in X$ such that

$$\begin{aligned}
&\text{(a)}\quad u_0(x_0) \geqslant u_0(r^1/n),\\
&\qquad\ u_1(x_1) \geqslant u_1(r^1/n, r^2/n),\\
&\qquad\ \ldots\\
&\qquad\ u_t(x_t) \geqslant u_t(r^t/n, r^{t+1}/n),\\
&\qquad\ \ldots\\
&\text{(b)}\quad \sum_t x_t^i \leqslant r^i, \qquad i = 1, 2, \ldots, \text{ and}
\end{aligned}$$

(c) the allocation $(x_0, x_1, \ldots, x_{t-1}, x_t^t)$ is Pareto optimal with respect to the utility functions $u_0(\cdot)$, $u_1(\cdot, \cdot), \ldots, u_{t-1}(\cdot, \cdot)$, $u_t(\cdot, x_t^{t+1})$.

We also need the set $\mathfrak{X}_n$ consisting of the elements $x \in X$ which satisfy (a) and (b). Clearly, $\mathfrak{X}_n$ is convex and compact.

Let $\mathfrak{U}_n(t-1)$ denote the subset of $\mathbb{R}^t$ which is defined as the image of $\mathfrak{X}_n$ by the mapping $x \mapsto u_0(x_0), \ldots, u_{t-1}(x_{t-1})$. Since $\mathfrak{X}_n$ is convex and hence arc-connected, $\mathfrak{U}_n(t-1)$ is arc-connected as the image of an arc-connected set by a continuous mapping.

Let $\mathfrak{P}_n(t, x_t^{t+1})$ denote the t-truncation of $\mathfrak{P}_n(t)$ associated with a *fixed* x_t^{t+1}, i.e., $\mathfrak{P}_n(t, x_t^{t+1}) = \{(x_0, x_1, \ldots, x_t^t)$, which are PO with respect to $u_0(\cdot)$, $u_1(\cdot), \ldots, u_{t-1}(\cdot)$, $u_t(\cdot, x_t^{t+1})\}$.

A.1.4. **Lemma.** *The set* $\mathfrak{P}_n(t, x_t^{t+1})$ *is homeomorphic to* $\mathfrak{U}_n(t-1)$, *every element of* $\mathfrak{P}_n(t, x_t^{t+1})$ *being continuously parameterized by the utility levels* $u_0, u_1, \ldots, u_{t-1}$.

Proof of (A.1.4). The lemma is a simple extension of known results; (cf. Arrow and Hahn [1, Chap. 5, Sect. 2, pp. 111–114] and Balasko [3, Appendix 3]). In the overlapping-generations model, there is no commodity which is an argument of every utility function, but the nature of generational overlap provides sufficient "relatedness" among consumers to prove the lemma. ∎

A.1.5. **Corollary.** *The set* $\mathfrak{P}_n(t, x_t^{t+1})$ *is arc-connected.*

Proof of (A.1.5). Obvious from (A.1.4) and the arc-connectedness of $\mathfrak{U}_n(t-1)$. ▮

Step 2. We now establish

A.1.6. LEMMA. *The set $\mathfrak{P}_n(t)$ is compact and arc-connected.*

Proof of (A.1.6). It follows from the inequalities $u_t(x_t) \geqslant u_t(r^t/n, r^{t+1}/n)$ that $\mathfrak{P}_n(t)$ is compact as a closed subset of the compact set X_n.

We must show that given x and x' in $\mathfrak{P}_n(t)$ they can be linked by a continuous path in $\mathfrak{P}_n(t)$. Fix x_t^{t+1}, x_{t+1}, x_{t+2},.... From (A.1.1), there is $x'' \in \mathfrak{P}_n(t, x_t^{t+1})$ such that

$$u_0(x_0'') = u_0(x_0'), \qquad u_1(x_1'') = u_1(x_1'),\dots, u_{t-1}(x_{t-1}'') = u_{t-1}(x_{t-1}').$$

From (A.1.5), x and x'' can be linked by a continuous path in $\mathfrak{P}_n(t, x_t^{t+1})$, hence in $\mathfrak{P}_n(t)$. Now take, for example, the segment linking $(x_t^{t+1}, x_{t+1},\dots)$ to $(x_t'^{t+1}, x_{t+1}',\dots)$. For any $\bar{x}_t^{t+1}$ belonging to the segment $[x_t^{t+1}, x_t'^{t+1}]$, there is by Lemma (A.1.4) a unique $x''(\bar{x}_t^{t+1}) \in \mathfrak{P}_n(t, \bar{x}_t^{t+1})$ such that $u_0(x_0''(\bar{x}_t^{t+1})) = u_0(x_0')$, $u_1(x_1''(\bar{x}_t^{t+1})) = u_1(x_1'),\dots,$ $u_{t-1}(x_{t-1}''(\bar{x}_t^{t+1})) = u_{t-1}(x_{t-1}')$. Furthermore, $x''(\bar{x}_t^{t+1})$ is a continuous function of $\bar{x}_t^{t+1}$. This construction defines a continuous path linking x'' to x' in $\mathfrak{P}_n(t)$. ▮

Step 3. Let $\mathfrak{P}(t)$ be the set of allocations $x \in X$ such that $(x_0, x_1,\dots, x_t^t)$ is Pareto-optimal with respect to the utility functions $u_0(\cdot)$, $u_1(\cdot),\dots,$ $u_{t-1}(\cdot)$, $u_t(\cdot, x_t^t)$. Note that there are now no restrictions on the utility levels $u_t(x_t)$. Then, we have

A.1.7. LEMMA. $\mathfrak{P}(t) = \bigcup_n \mathfrak{P}_n(t)$.

Proof. Obvious. ▮

A.1.8. LEMMA. $\mathfrak{P} = \bigcap_t \mathfrak{P}(t)$.

Proof. Straightforward. (A similar construction is used in the Balasko and Shell [5] study of competitive equilibrium; cf. [5, Proposition (3.10) and Remark (3.12)].) ▮

Step 4. From (A.1.7) and (A.1.8), we have

$$\mathfrak{P} = \left(\bigcap_t \bigcup_n \mathfrak{P}_n(t)\right)$$

and hence

$$\mathfrak{P} = \left(\bigcup_n \bigcap_t \mathfrak{P}_n(t)\right).$$

Clearly, we have

$$\cdots \subset \mathfrak{P}_n(t+1) \subset \mathfrak{P}_n(t) \subset \cdots.$$

Therefore, $\bigcap \mathfrak{P}_n(t)$ is the intersection of a decreasing sequence of compact, connected, nonempty sets, hence is compact and connected. (Cf. e.g., Bourbaki [8, Chap. II, Sect. 4, Exercise 14, p. 212]. For a proof in a metric setting, cf., e.g., Lefschetz [14, Chap. I, Sect. 4, pp. 7–9].)

The sequence $\{\bigcap_t \mathfrak{P}_n(t)\}$ is clearly increasing in n, hence $\mathfrak{P} = \bigcup_n (\bigcap_t \mathfrak{P}_n(t))$ is connected; cf., e.g., Bourbaki [18, Chap. I, Sect 11, Proposition 2, p. 108]. ▮

A.1.9. *Remark.* We have shown only that $\mathfrak{P}$ is connected although we believe that one can probably establish the stronger result of arc-connectedness. (Arc-connectedness of $\mathfrak{P}$ becomes equivalent here to local arc-connectedness. It is not, however, in general true that the intersection of a nested sequence of compact, arc-connected, nonempty sets is arc-connected.)

A.2. *The Set of Pareto-Optima*

Let $\mathfrak{P}$ denote the set of Pareto-optimal allocations. Let $\mathfrak{U}_1$ be the image of X under the continuous mapping

$$x \longmapsto (u_1(x_1),\ldots, u_t(x_t),\ldots).$$

Clearly, $\mathfrak{U}_1$ is arc-connected. Next we show that $\mathfrak{P}$ is homeomorphic to $\mathfrak{U}_1$ and hence $\mathfrak{P}$ is arc-connected.

A.2.1. LEMMA. *The set of Pareto-optimal allocations $\mathfrak{P}$ is homeomorphic to $\mathfrak{U}_1$.*

Proof of (A.2.1). Essentially, we repeat the proof used for analyzing the structure of the PO set for finite pure-exchange economies. Cf. (A.1.4) above and Balasko [2, Corollary 1, p. 564].

Let $(\bar{u}_1, \bar{u}_2,\ldots, \bar{u}_t,\ldots) \in \mathfrak{U}_1$ be fixed. Consider the set $\{x \in X \mid u_1(x_1) = \bar{u}_1,\ldots, u_t(x_t) = \bar{u}_t,\ldots\}$, which is nonempty because $(\bar{u}_1,\ldots, \bar{u}_t,\ldots)$ belongs to $\mathfrak{U}_1$. Furthermore, the set is clearly closed and every component x_t is bounded; therefore, the set is compact. The utility function u_0 can be viewed as a continuous function of the sequence x. Hence, it has a maximum on the compact set $\{x \in X \mid u_1(x_1) = \bar{u}_1,\ldots, u_t(x_t) = \bar{u}_t,\ldots\}$. By the strict quasiconcavity of every $u_t(\cdot)$, this maximum is unique and is a continuous function of $(\bar{u}_1,\ldots, \bar{u}_t,\ldots)$. We have thus defined a continuous mapping from $\mathfrak{U}_1$ into $\mathfrak{P}$. The inverse mapping $x \in \mathfrak{P} \longmapsto (u_1(x_1),\ldots, u_t(x_t),\ldots)$ is obviously continuous.

A.2.2. **Proposition.** *The set of Pareto-optimal allocations $\mathfrak{P}$ is arc-connected.*

Proof. Follows directly from (A.2.1) and the arc-connectedness of $\mathfrak{U}_1$. ∎

Glossary of Basic Notation

$x = (x_0, x_1, \ldots)$ is the (intertemporal) commodity allocation sequence in the allocation space $X = \mathbb{R}^l_{++} \times \mathbb{R}^{2l}_{++} \times \mathbb{R}^{2l}_{++} \times \cdots$.

$\omega = (\omega_0, \omega_1, \ldots) \in X$ is the sequence of commodity endowments and $r = \sum_t \omega_t$ is the sequence of resources (or, aggregate endowments).

$p = (p^1, p^2, \ldots) \in \mathbb{R}^l_{++} \times \mathbb{R}^l_{++} \times \cdots$ is the (intertemporal) commodity price sequence. $\mathscr{P}$ is the set of normalized prices $\{p \mid p^{1,1} = 1\}$, $\mathscr{P} \subset \mathbb{R}^l_{++} \times \mathbb{R}^l_{++} \times \cdots$.

The monetary policy (or, monetary transfer) sequence is $m = (m^1_0, m^1_1, m^2_1, m^2_2, m^3_2, \ldots) \in M = \mathbb{R}^\infty$, the set of such monetary policies (or, monetary transfers). We also refer to $\mu = (\mu_0, \mu_1, \mu_2, \ldots) = (m^1_0, m^1_1 + m^2_1, m^2_2 + m^3_2, \ldots)$ as a monetary policy. The set of such μ is $\mathscr{M} = \mathbb{R}^\infty$.

$m^t \in \mathbb{R}$ is the aggregate money supply at time $t \geqslant 1$.

$x^{s,m}_t \in \mathbb{R}$ is the gross addition to money balances committed by consumer t in period s; x^m is the sequence $(x^{1,m}_0, x^{1,m}_1, x^{2,m}_1, \ldots)$.

$p^m \in \mathbb{R}_+$ is the present price of money.

$q = (p, p^m) \in \mathscr{Q} = \{(p, p^m) \mid p \in \mathscr{P}$ and $p^m \in \mathbb{R}_+\} \subset \mathbb{R}_+ \times \mathbb{R}^l_{++} \times \mathbb{R}^l_{++} \times \cdots$.

$\mathscr{Q}(\omega, \mu) \subset \mathscr{Q}$ is the set of monetary competitive equilibria $q = (p, p^m)$ associated with the endowments $\omega \in X$ and the monetary policy $\mu \in \mathscr{M}$.

$\mathscr{P}^m(\omega, \mu) = \{p^m \mid (p, p^m) \in \mathscr{Q}(\omega, \mu)\} \subset \mathbb{R}^+$ is the set of equilibrium money prices given the endowment $\omega \in X$ and the monetary policy $\mu \in \mathscr{M}$.

$w = (w_0, w_1, \ldots)$ is the sequence of "incomes" in the set $W = \mathbb{R}^\infty_{++}$.

$\mathscr{B} \subset \mathscr{P} \times W$ is the set of price-income equilibria (p, w) given fixed resources r.

$\mathscr{M}_B(\omega) \subset \mathscr{M}$ is the set of normalized ω-bonafide monetary policies. The set of ω-bonafide monetary policies is the cone in $\mathscr{M}$ generated by $\mathscr{M}_B(\omega)$. $\mathscr{M}_B \subset \mathscr{M}$ is the set of normalized bonafide monetary policies. The set of bonafide monetary policies is the cone in $\mathscr{M}$ generated by $\mathscr{M}_B$. $\mathscr{M}_P(\omega)$ is the set of potent ω-monetary policies.

$\bar{\mathfrak{P}} \subset X$ is the set of weakly-Pareto-optimal (WPO) allocations. $\mathfrak{P}$ is the set of Pareto-optimal (PO) allocations, thus $\mathfrak{P} \subset \bar{\mathfrak{P}}$.

ACKNOWLEDGMENTS

We thank Dave Cass and John Bigelow for their critical advice and very helpful comments.

REFERENCES

1. K. J. ARROW AND F. H. HAHN, "General Competitive Analysis," Holden-Day, San Francisco, 1971.
2. Y. BALASKO, Economic equilibrium and catastrophe theory, *Econometrica* **46** (1978), 557–569.
3. Y. BALASKO, Budget-constrained and Pareto-efficient allocations, *J. Econ. Theory* **21** (1979), 359–379.
4. Y. BALASKO, D. CASS, AND K. SHELL, Existence of competitive equilibrium in a general overlapping-generations model, *J. Econ. Theory* **23** (1980), 307–322.
5. Y. BALASKO AND K. SHELL, The overlapping-generations model, I: The case of pure exchange without money, *J. Econ. Theory* **23** (1980), 281–306.
6. Y. BALASKO AND K. SHELL, The overlapping-generations model, III: the log-linear case, *J. Econ. Theory* **24** (1981), 143–152.
7. Y. BALASKO AND K. SHELL, "On Taxation and Competitive Equilibria," CARESS Working Paper 80–18, University of Pennsylvania, Philadelphia, 1980.
8. N. BOURBAKI, "Elements of Mathematics: General Topology, Part I," Addison-Wesley, Reading, Mass., 1966.
9. D. CASS, M. OKUNO, AND I. ZILCHA, The role of money in supporting the Pareto-optimality of competitive equilibrium in consumption-loan type models, *J. Econ. Theory* **20** (1979), 41–80.
10. D. CASS AND K. SHELL, In defense of a basic approach, *in* "Models of Monetary Economies" (J. Kareken and N. Wallace, Eds.), pp. 251–203, Federal Reserve Bank of Minneapolis, 1980.
11. D. CASS AND M. E. YAARI, A re-examination of the pure consumption loans model, *J. Pol. Econ.* **74** (1966), 353–367.
12. M. FRIEDMAN, "Essays in Positive Economics," Univ. of Chicago Press, Chicago, 1953.
13. D. GALE, Pure exchange equilibrium of dynamic economic models, *J. Econ. Theory* **6** (1973), 12–36.
14. S. LEFSCHETZ, "Differential Equations: Geometric Theory," Second ed., Interscience, New York, 1962.
15. M. OKUNO AND I. ZILCHA, On the efficiency of competitive equilibrium in infinite horizon monetary models, unpublished manuscript, 1978.
16. D. PATINKIN, "Money, Interest, and Prices," Row, Peterson, Evanston, Ill., 1955.
17. P. A. SAMUELSON, An exact consumption-loan model of interest with or without the social contrivance of money, *J. Pol. Econ.* **66** (1958), 467–482.
18. K. SHELL, Notes on the economics of infinity, *J. Pol. Econ.* **79** (1971), 1002–1011.
19. N. WALLACE, The overlapping-generations model of fiat money, *in* "Models of Monetary Economies" (J. Kareken and N. Wallace, Eds.), pp. 49–78, Federal Reserve Bank of Minneapolis, 1980.

JOURNAL OF ECONOMIC THEORY 24, 143–152 (1981)

The Overlapping-Generations Model. III. The Case of Log-Linear Utility Functions*

YVES BALASKO

Université Paris I, CEPREMAP, Paris, France,
and Center for Analytic Research in Economics and the Social Sciences,
University of Pennsylvania, Philadelphia, Pennsylvania 19104

AND

KARL SHELL

Center for Analytic Resarch in Economics and the Social Sciences,
University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received December 14, 1979; revised August 12, 1980

1. INTRODUCTION

Existence of competitive equilibria in the overlapping-generations model has been studied in some detail; see [4–6, 9]. We need to know more about the properties of these equilibria. How many are there? How do they depend on the basic parameters of the economy? We also need to compare the set of monetary competitive equilibria with the set of nonmonetary competitive equilibria. Are there "vastly more" monetary equilibria than nonmonetary equilibria?

Analysis of the properties of competitive equilibria is for the overlapping-generations model an important, but difficult, task. In his seminal article [8], Samuelson put it this way: "... if we take any finite stretch of time and write out the equilibrium conditions, we always find them containing discount rates from before the finite period and discount rates from afterward. We never seem to get enough equations; lengthening our time periods turns out always to add as many new unknowns as it supplies equations ..." Samuelson suggested that, rather than facing this difficulty directly, "We can try to cut the Gordian knot by our special assumption of stationariness ..."

---

* This research is part of a larger project on the overlapping-generations model and macroeconomics, joint with Dave Cass and supported by Grants SOC-78-06157 and SES-80-07012 from the National Science Foundation to the Center for Analytic Research in Economics and the Social Sciences at the University of Pennsylvania.

0022-0531/81/010143-10$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved.

Indeed, assuming stationarity is just that—*cutting* the Gordian knot. An important goal of government macroeconomic policy is stabilization of the short-run fluctuations in an evolving and changing economic environment. All this is lost when we focus solely on the long-run steady state of a replicating economy. One of the defects of recent monetary theory has been its emphasis on the long run, a very peculiar point of view for the stabilization-oriented macroeconomist.

At any rate, stationary analyses are not appropriate to the overlapping-generations model which we considered in [5, 6], where preferences and endowments are allowed to vary from consumer to consumer in a general way. It turns out, however, for the special (nonstationary) case where utility functions are log-linear, the Gordian knot is easily *untied*. Log-linearity drastically simplifies the mathematical analysis and reduces it to a study of the properties of an infinite number of linear equations having an infinite number of unknowns. The structure of the overlapping-generations model then enables us to solve this system in a straightforward and elementary way by merely considering finite stretches of time.

Section 2 is devoted to the log-linear nonmonetary economy. The unique Walrasian equilibrium is shown to be a linear function (i.e., a degree-one polynomial function, with a "constant" term) of endowments, considered as the parameters of the economy. The log-linear monetary economy is studied in Section 3. There is a nonnegative scalar $\bar{p}^m$ with the property that if the price of money p^m exceeds $\bar{p}^m$ then there is no competitive equilibrium, and if $p^m \in [0, \bar{p}^m)$ there is exactly one competitive equilibrium. This equilibrium is a linear function of p^m. Note that in general the price of money is indeterminate and parameterizes the monetary equilibria. This result confirms our experience with monetary overlapping-generations models: we frequently encounter a vast multiplicity of equilibria.

The analysis of the nonmonetary and monetary economies is extended in Section 4 to the case where only consumer t and his successors possess log-linear utility functions. Here, too, the Gordian knot can be untied: equilibrium prices are determined by a counting-equations approach, providing a generic solution to the problem of the number of equilibria and their behavior with respect to the parameters defining the economy. Extension of our results beyond the "nearly log-linear" case seems to require the development of a transversality theory which can be applied to the appropriate infinite-dimensional spaces. How to untie this infinite-dimensional "Gordian knot" remains for the general case an unsolved problem, which is important and challenging.

2. THE LOG-LINEAR NONMONETARY ECONOMY

We use the notation of Balasko and Shell [5]. Let

$$u_0(x_0)=\sum_{k=1}^{l} a_0^{1,k}\log x_0^{1,k}$$

and

$$u_t(x_t)=\sum_{k=1}^{l} a_t^{t,k}\log x_t^{t,k}+\sum_{k=1}^{l} a_t^{t+1,k}\log x_t^{t+1,k} \qquad \text{for} \quad t\geqslant 1,$$

be the utility functions of consumer 0 and consumer t, respectively. The coefficients $(a_t^{s,k})$ are assumed to satisfy the restrictions

$$a_0^{1,k}>0, \qquad a_t^{t,k}>0, \qquad a_t^{t+1,k}>0,$$
$$\sum_{k=1}^{l} a_0^{1,k}=1, \qquad \sum_{k=1}^{l} a_t^{t,k}+\sum_{k=1}^{l} a_t^{t+1,k}=1.$$

Recall that the normalized price sequence $p\in\mathcal{P}=\{p\mid p^{1,1}=1\}$ is a Walrasian equilibrium (cf. [5, Definition (2.3)]) associated with the endowment sequence $\omega=(\omega_0,\omega_1,\ldots,\omega_t,\ldots)\in X$ if and only if

$$\sum_t f_t(p, p\cdot\omega_t)=\sum_t \omega_t=r.$$

This equation is equivalent to the equation system

$$\sum_t f_t(p, w_t)=\sum_t \omega_t=r,$$
$$p\cdot\omega_t=w_t \qquad \text{for} \quad t\geqslant 0,$$

where the unknowns are not only prices $p\in\mathcal{P}$ but also "incomes" $w=(w_0,w_1,\ldots,w_t,\ldots)\in W$. If we substitute for the demand function f_t the particular expression resulting from the constrained maximization of the log-linear utility function $u_t(\cdot)$, we obtain the system of equations

$$a_{t-1}^{t,i}w^{t-1}+a_t^{t,i}w_t=p^{t,i}\cdot r^{t,i} \qquad \text{for} \quad t\geqslant 1 \text{ and } i=1,\ldots,l,$$
$$w_t=p\cdot\omega_t \qquad \text{for} \quad t\geqslant 0,$$

with $p^{t+1}>0$ and $w_t>0$ for $t\geqslant 0$. A price sequence $p\in\mathcal{P}$ which solves the above system is a Walrasian equilibrium associated with ω and conversely.

2.1. **Proposition.** *There is a unique Walrasian equilibrium $p \in \mathscr{P}$ (and corresponding allocation $x \in X$) associated with every endowment sequence $\omega \in X$.*

Proof. Consider the truncated equilibrium system

$$
\begin{aligned}
&a_0^{1,1} w_0 + a_1^{1,1} w_1 = r^{1,1}, \\
&\quad \cdot \quad \cdot \quad \cdot \\
&a_t^{t+1,l} w_t + a_{t+1}^{t+1,l} w_{t+1} = p^{t+1,l} \cdot r^{t+1,l}, \\
&w_0 = p^1 \cdot \omega_0^1, \\
&w_1 = p^1 \cdot \omega_1^1 + p^2 \cdot \omega_1^2, \\
&\quad \cdot \quad \cdot \quad \cdot \\
&w_t = p^t \cdot \omega_t^t + p^{t+1} \cdot \omega_t^{t+1}, \\
&p^1 > 0, \dots, p^{t+1} > 0, \\
&w_0 > 0, \dots, w_t > 0.
\end{aligned}
\tag{2.1.t}
$$

Clearly, if $p \in \mathscr{P}$ is a Walrasian equilibrium associated with ω, then the truncated vector $(p^1, \dots, p^{t+1}, w_0 = p \cdot \omega_0, \dots, w_{t+1} = p \cdot \omega_{t+1})$ is a solution to (2.1.t). From Balasko and Shell [5, Proposition (3.10)], there is at least one Walrasian equilibrium $p \in \mathscr{P}$; therefore, by truncation, there is at least one solution of the system (2.1.t) for each $t \geqslant 0$. We next show that the solution to (2.1.t) is unique.

Assume that there are at least two distinct solutions, $\pi = (p^1, \dots, p^{t+1}, w_0, \dots, w_{t+1})$ and $\pi' = (p'^1, \dots, p'^{t+1}, w'_0, \dots, w'_{t+1})$. We claim that there is then $s \in \{1, \dots, t+1\}$ such that $p^s \neq p'^s$ or $s \in \{0, \dots, t\}$ such that $w_s \neq w'_s$; i.e., we claim that, if π and π' are distinct solutions, they do not differ solely in the last components, w_{t+1} and w'_{t+1}. To prove the claim, assume that $p^s = p'^s$ for $s = 1, \dots, t+1$ and $w_s = w'_s$ for $s = 0, \dots, t$. Select any $i \in \{1, \dots, l\}$, then

$$a_t^{t+1,l} w_t + a_{t+1}^{t+1,i} w_{t+1} = p^{t+1,l} \cdot r^{t+1,l}$$

and

$$a_t^{t+1,l} w'_t + a_{t+1}^{t+1,l} w'_{t+1} = p'^{t+1,l} \cdot r^{t+1,l},$$

which together imply that $a_{t+1}^{t+1,l}(w_{t+1} - w'_{t+1}) = 0$, or $w_{t+1} = w'_{t+1}$, implying in turn that $\pi = \pi'$, contrary to assumption. Thus, if π and π' are distinct solutions, they cannot differ solely in their last components.

Let $\bar{\pi} = \lambda \pi + (1 - \lambda) \pi'$, where $\lambda \in \mathbb{R}$. Clearly, if $\lambda \notin [0, 1]$, then $\bar{\pi}$ is a solution to the equalities of (2.1.t), which might violate the inequalities of (2.1.t). Recall that $\bar{\pi}$ is linear in $\lambda \in \mathbb{R}$ and that (from the previous result) at least one of the first $(l+1)(t+1)$ components of $\bar{\pi}$ is not constant in λ. Therefore, there is a $\lambda \in \mathbb{R}$ such that the vector $(\bar{p}^1,\dots,\bar{p}^{t+1}, \bar{w}_0\dots,\bar{w}_t) = \lambda(p^1,\dots,p^{t+1},w_0,\dots,w_t)+(1-\lambda)(p'^1,\dots,p'^{t+1},w'_0,\dots,w'_t)$ is nonnegative but equal to zero in at least one component. Therefore, either for some $k \in \{1,\dots,l\}$ and some $s \in \{1,\dots,t+1\}$, $\bar{p}^{s,k}=0$ or for some $s \in \{0,\dots,t\}$, $\bar{w}_s = 0$. Then, from the equations in (2.1.t), either $\bar{w}_s = \bar{w}_{s-1} = 0$ or $\bar{p}^s = \bar{p}^{s+1} = 0$, which implies that $(\bar{p}^1,\dots,\bar{p}^{t+1}, \bar{w}_0,\dots,\bar{w}_t) = 0$. This contradicts the assumption that $a_0^{1,1}w_0 + a_1^{1,1}w_1 = r^{1,1} > 0$. The rest of the proof follows by varying t. ∎

Proposition (2.1) extends a well-known result on the uniqueness of Walrasian equilibrium from the standard finite model to the overlapping-generations model; see, e.g., [1].

3. THE LOG-LINEAR MONETARY ECONOMY

We use the notation of Balasko and Shell [6]. Let $p^m \in \mathbb{R}_+$ be the present price of money and let $\mu = (\mu_0, \mu_1,\dots,\mu_t,\dots) \in \mathscr{M}$ be the monetary policy. The equilibrium conditions are then

$$\begin{aligned} & a_{t-1}^{t,i}w_{t-1} + a_t^{t,i}w_t = p^{t,i}r^{t,i} && \text{for } t \geqslant 1 \text{ and } i = 1,\dots,l, \\ & w_t = p\cdot\omega_t + p^m\mu_t && \text{for } t \geqslant 0, \\ & p^{t+1} > 0, \quad w_t > 0 && \text{for } t \geqslant 0. \end{aligned}$$

3.1. PROPOSITION. *The set $\mathscr{M}_B(\omega)$ of ω-bonafide normalized monetary policies is convex.*

Proof. The set of price-income equilibria (p, w) (cf. [6, Definition (5.4)]) is defined by a set of linear equations and linear inequalities and hence is obviously convex. The set $\mathscr{M}_B(\omega)$ is the image of the set of price-income equilibria by the linear mapping Φ_ω: $(p, w_0,\dots,w_t,\dots) \mapsto (w_0 - p\cdot\omega_0,\dots,w_t - p\cdot\omega_t,\dots)$ which maps convex sets onto convex sets (cf. [6, Section 6, especially Proposition (6.2)]). Therefore, $\mathscr{M}_B(\omega)$ is convex.

3.2. PROPOSITION. *The set of equilibrium present prices of money $\mathscr{P}^m(\omega,\mu)$ is an interval for fixed endowments $\omega \in X$ and fixed monetary policy $\mu \in \mathscr{M}$.*

Proof. There is a linear bijection between $\mathscr{P}^m(\omega,\mu)$ and the intersection of the set $\mathscr{M}_B(\omega)$ with the ray $L(\mu)$; cf. Balasko and Shell [6, Proposition 7.3]. Since $\mathscr{M}_B(\omega)$ and $L(\mu)$ are each convex sets, their intersection, $\mathscr{M}_B(\omega) \cap L(\mu)$, is also convex. Therefore, $\mathscr{P}^m(\omega,\mu) \subset \mathbb{R}_+$ is convex and thus an interval. ∎

Hence, there exists some nonnegative scalar $\bar{p}^m$ (possibly equal to $+\infty$; see, however, Balasko and Shell [6, Proposition (7.4)]) such that $\mathscr{P}^m(\omega, \mu)$ is the interval with bounds 0 and $\bar{p}^m$. It may happen that $\mathscr{M}_B(\omega) \cap L(\mu) = \{0\}$, in which case $\bar{p}^m = 0$.

3.3. Proposition. *For any $p^m \in [0, \bar{p}^m)$ there is a unique monetary competitive equilibrium $q = (p, p^m) \in \mathscr{Q}(\omega, \mu)$ associated with (ω, μ).*

Proof. From Proposition (3.2) and the definition of $\bar{p}^m$, it follows that for any $p^m \in [0, \bar{p}^m)$, the following system has at least one solution

$$\begin{aligned}
&a_{t-1}^{t,i} w_{t-1} + a_t^{t,i} w_t = p^{t,i} r^{t,i} && \text{for} \quad t \geqslant 1 \text{ and } i = 1, \dots, l, \\
&w_t = p \cdot \omega_t + p^m \mu_t && \text{for} \quad t \geqslant 0, \\
&p^{t+1} > 0, \qquad w_t > 0 && \text{for} \quad t \geqslant 0.
\end{aligned} \tag{3.3.1}$$

The following truncated system admits as a solution the truncation of any solution of (3.3.1)

$$\begin{aligned}
&a_0^{1,1} w_0 + a_1^{1,1} w_1 = r^{1,1}, \\
&\quad \cdots \\
&a_t^{t+1,l} w_t + a_{t+1}^{t+1,l} w_{t+1} = p^{t+1,l} r^{t+1,l}, \\
&w_0 = p^1 \cdot \omega_0 + p^m \mu_0, \\
&\quad \cdots \\
&w_t = p^t \cdot \omega_t^t + p^{t+1} \cdot \omega_t^{t+1} + p^m \mu_t, \\
&p^1 > 0, \dots, p^{t+1} > 0, \\
&w_0 > 0, \dots, w_t > 0.
\end{aligned} \tag{3.3.t}$$

Notice that the equations in (3.3.t) have the same coefficients as the equations in (2.1.t); only the "constant terms" differ. Uniqueness of the solution to (2.1.t) implies that the equations in (2.1.t) define a nondegenerate system. Since nondegeneracy is independent of the values of the "constant terms," the linear equations in (3.3.t) are nondegenerate and have a unique solution. The proof is completed as in Proposition (2.1). ▮

3.4. Proposition. *The monetary competitive equilibrium $q = (p, p^m) \in \mathscr{Q}(\omega, \mu)$ is a linear function of p^m for each $p^m \in [0, \bar{p}^m)$.*

Proof. Since we merely need to establish this property for each coordinate of p, it suffices to analyze the truncated system (3.3.t). The proposition is then obvious after applying the explicit Cramer formula used in solving a nondegenerate system of linear equations. ▮

4. ECONOMIES IN WHICH ALL BUT A FINITE NUMBER OF UTILITY FUNCTIONS ARE LOG-LINEAR

Assume that for $t' \geqslant t$, the utility function $u_{t'}(\cdot)$ is log-linear, i.e., consumer t and his followers have log-linear functions defined on $\mathbb{R}^{2l}_{++}$. Given this assumption—that the economy is "nearly log-linear"—we are able in what follows to answer completely questions about the number and behavior of equilibria.

Consider first the nonmonetary case. Let $\omega \in X$ be the sequence of endowments. We hold total resources, $r = \sum_t \omega_t$, constant and study how the equilibria vary as ω varies.

Define the t-truncation $\omega(t)$ of ω as the vector

$$\omega(t) = (\omega_0^1, \omega_1^1, \omega_1^2, ..., \omega_{t-1}^{t-1}, \omega_{t-1}^t, \omega_t^t).$$

Notice that this truncated vector is not the same as the one used in Balasko and Shell [5, Lemma (3.9) and Proposition (3.10)]. Let $\mathfrak{X}_t = \mathbb{R}^{2lt}_{++}$ be the set of all truncated allocations $x(t) = (x_0^1, x_1^1, x_1^2, ..., x_{t-1}^{t-1}, x_{t-1}^t, x_t^t)$.

4.1. **PROPOSITION.** *There is an open dense subset $\mathfrak{R}_t$ of $\mathfrak{X}_t$ such that if $\omega(t)$, the truncation of $\omega \in X$, belongs to $\mathfrak{R}_t$, then*:

(a) *There is only a finite ($\neq 0$) number of nonmonetary equilibria $p \in \mathscr{P}$ associated with the endowment sequence ω*;

(b) *There is a neighborhood $\mathfrak{V} \subset \mathfrak{R}_t$ such that for any nonmonetary equilibrium $p \in \mathscr{P}(\omega)$ the prices $p^1, p^2, ..., p^t$ are smooth functions of the endowments $\omega(t) \in \mathfrak{V}$.*

Proof. We decompose the system of equilibrium equations in the following way:

$$\left.\begin{aligned}
&f_0^1(p^1, w_0) + f_1^1(p^1, p^2, w_1) = r^1,\\
&\quad\cdot\quad\cdot\quad\cdot\\
&f_{t-2}^{t-1}(p^{t-2}, p^{t-1}, w_{t-2}) + f_{t-1}^{t-1}(p^{t-1}, p^t, w_{t-1}) = r^{t-1},\\
&f_{t-1}^{t,1}(p^{t-1}, p^t, w_{t-1}) + a_t^{t,1}\frac{w_t}{p^{t,1}} = r^{t,1},\\
&\quad\cdot\quad\cdot\quad\cdot\\
&f_{t-1}^{t,l}(p^{t-1}, p^t, w_{t-1}) + a_t^{t,l}\frac{w_t}{p^{t,l}} = r^{t,l},\\
&w_0 = p\cdot\omega_0, w_1 = p\cdot\omega_1, ..., w_{t-1} = p\cdot\omega_{t-1},\\
&p^1 > 0, ..., p^t > 0,
\end{aligned}\right\} \qquad (4.1.1)$$

$$\left.\begin{aligned}
&a_t^{t+1,1}\frac{w_t}{p^{t+1,1}} + a_{t+1}^{t+1,1}\frac{w_{t+1}}{p^{t+1,1}} = r^{t+1,1},\\
&\cdot\quad\cdot\quad\cdot\\
&\cdot\quad\cdot\quad\cdot\\
&w_t = p\cdot\omega_t,\ w_{t+1} = p\cdot\omega_{t+1},\dots,\\
&p^{t+1} > 0,\ p^{t+2} > 0,\dots.
\end{aligned}\right\} \qquad (4.1.2)$$

In System (4.1.1), there are exactly as many unknowns ($p^1, p^2,\dots, p^t, w_0,\dots, w_{t-1}, w_t$) as there are equations. Assuming that this implies determination of the solutions of (4.1.1), a subject which we shall reconsider in a moment, we can then turn to (4.1.2), where now w_t and p^t are given and where the unknowns are w_{t+1}, $w_{t+2},\dots, p^{t+1}$, $p^{t+2},\dots$ In fact, the equation system (4.1.2) is equivalent to a linear equation system which has basically the same structure as the linear equation system considered in Proposition (2.1) and hence its structure can be investigated by the same truncation process as in (2.1.t). Therefore, once w_t and p^t are known, the equalities in (4.1.2) have a unique solution which does or does not satisfy the sign constraints. This unique solution is a linear function of $\omega_t, \omega_{t+1},\dots$ and of p^t and w_t.

Let us return to the counting-equations argument for System (4.1.1). In order to preserve the elementary nature of this paper, we merely sketch the way in which that argument can be employed to establish the generic finiteness of the number of solutions of (4.1.1). This argument parallels the one used in the standard general-equilibrium model by Balasko [2], an article to which the reader is referred for working out the details of the proof. One considers in the space of ($p^1,\dots, p^t$, $w_0,\dots, w_t$) the subset defined by the equations

$$f_{i-1}^i(p, w_{i-1}) + f_i^i(p, w_i) = r^i \qquad \text{for} \quad i \geqslant 1.$$

Using, for example, the regular value theorem, one shows that this set is a smooth submanifold of the set of prices and incomes. We call it the section manifold. Equilibria correspond to the intersection of the section manifold with the linear manifold defined by the equations

$$w_0 = p\cdot\omega_0,\dots, w_{t-1} = p\cdot\omega_{t-1}.$$

Transversality of these two manifolds and their dimensions imply discreteness of the intersection set. The boundary behavior of individual demand functions implies that this intersection is compact, hence finite. (In the Appendix to Balasko [3], it is shown that the section manifold is properly embedded with respect to the set of linear manifolds defined by equations of the type $w_0 = p \cdot \omega_0, \ldots, w_{t-1} = p \cdot \omega_{t-1}$. This is exactly what is needed here.) We conclude by an application of Thom's transversality theorem; see [7, Theorem (4.9)].

The second part of Proposition (4.1) is an obvious consequence of the transversality properties. Note that the linear structure of (4.1.2) gives us a complete answer to the question of the local behavior of equilibria as a function of initial endowments. ∎

Let us now study the case of monetary equilibria. Let $\mu \in \mathscr{M}$ be a monetary policy. To ensure existence of proper monetary equilibria, we assume that $\mu \in \mathscr{M}_B(\omega)$, i.e., μ is a bonafide monetary policy. In view of the results of Section 3, where we have shown that monetary equilibria can be parameterized by the initial endowments ω, the monetary policy μ, and the price of money p^m, the next step is to study how the equilibrium set

$$\mathscr{P}(\omega, \mu, p^m) = \{p \mid (p, p^m) \in \mathscr{Q}(\omega, \mu)\} \subset \mathscr{P}$$

varies as (ω, μ, p^m) varies.

Let $\mu(t) = (\mu_0, \mu_1, \ldots, \mu_t)$ denote the t-truncation of t. Fixing $(\mu_{t+1}, \mu_{t+2}, \ldots)$ and $(\omega_t^{t+1}, \omega_{t+1}^{t+1}, \omega_{t+1}^{t+2}, \ldots)$ let $\mathfrak{E}$ be the subset of $\mathbb{R}^l_{++} \times \mathbb{R}^{2l}_{++} \cdots \times \mathbb{R}^{2l}_{++} \times \mathbb{R}^{l+1} \times \mathbb{R}_+$ generated by the $(\omega(t), \mu(t), p^m)$ that satisfy $\mu \in \mathscr{M}_B(\omega)$ and $\mathscr{P}(\omega, \mu, p^m) \neq \varnothing$.

4.2. Proposition. *There is an open dense subset* $\mathfrak{E}_t$ *of* $\mathfrak{E}$ *such that for any* $(\omega(t), \mu(t), p^m)$ *in* $\mathfrak{E}_t$, *then* $\mathscr{P}(\omega, \mu, p^m)$ *is a finite set.*

Proof. As in the case of Proposition (4.1), we decompose the equilibrium equations into two systems such that a solution of the t-truncated equation system determines in turn a solution of the second system which is in fact a linear equation system. Therefore, what is at stake once again is the t-truncation equation system. Proposition (4.2) merely states a transversality property of the linear space

$$w_0 = p \cdot \omega_0 + p^m \mu_0, \ldots, w_t = p \cdot \omega_t + p^m \mu_t$$

with the section manifold. This transversality property obviously holds for an open dense set of (ω, μ, p^m). ∎

Note that this proof also shows that when $(\omega(t), \mu(t), p^m)$ varies in $\mathfrak{E}_t$, then every equilibrium $p \in \mathscr{P}(\omega, \mu, p^m)$ can be considered as a smooth function coordinatewise of ω, μ, and p^m.

ACKNOWLEDGMENTS

We are grateful to Dave Cass and John Bigelow for their comments.

REFERENCES

1. Y. BALASKO, Some results on uniqueness and on stability in general equilibrium theory, *J. Math. Econ.* **2** (1975), 95–118.
2. Y. BALASKO, Economic equilibrium and catastrophe theory: an introduction, *Econometrica* **48** (1978), 557–569.
3. Y. BALASKO, Budget-constrained Pareto-efficient allocations, *J. Econ. Theory* **21** (1979), 359–379.
4. Y. BALASKO, D. CASS, AND K. SHELL, Existence of competitive equilibrium in a general overlapping-generations model, *J. Econ. Theory* **23** (1980), 307–322.
5. Y. BALASKO AND K. SHELL, The overlapping-generations model, I: The case of pure exchange without money, *J. Econ. Theory* **23** (1980), 281–306.
6. Y. BALASKO AND K. SHELL, The overlapping-generations model, II: The case of pure exchange with money, *J. Econ. Theory* **24** (1981), 112–142.
7. M. GOLUBITSKY AND V. GUILLEMIN, "Stable Mappings and their Singularities," Springer, New York, 1973.
8. P. A. SAMUELSON, An exact consumption-loan model of interest with or without the social contrivance of money, *J. Pol. Econ.* **66** (1958), 477–482.
9. C. A. WILSON, Equilibrium in dynamic models with an infinity of agents, *J. Econ. Theory* **24** (1981), 95–111.

JOURNAL OF ECONOMIC THEORY 24, 153–154 (1981)

Acknowledgment

The *Journal of Economic Theory* is indebted to many people for the editorial advice which they have given. We take this opportunity to thank the scholars who refereed manuscripts during 1980.

The Editorial Board

Beth Allen
Kenneth J. Arrow
A. B. Atkinson

Yves Balasko
Taradas Bandyopadhyay
Salvador Barberá
Peter Berck
Truman Bewley
Rabi Bhattacharya
Sudipto Bhattacharya
G. O. Bierwag
Charles Blackorby
Douglas H. Blair
Douglas T. Breedon
Jan Brueckner
Edwin Burmeister
Gerald R. Butters

Xavier Calsamiglia
Donald Campbell
Graciela Chichilnisky
John S. Chipman
Russell Cooper
Bernard Cornet
John Cox
Jacques Crémer
W. Davis Dechert
W. E. Diewert
David J. Donaldson
Pradeep Dubey
Bhaskar Dutta
Philip H. Dybvig

David Easley
Ivar Ekeland
Wilfred Ethier

Joseph Farrell
Allan M. Feldman
Thomas Finn
Franklin M. Fisher
Jeff Frankel
James W. Friedman
Gérard Fuchs
Carl Futia

W. Gaertner
Robert E. Gaines
Douglas Gale
Carlos Garcia
Peter Gärdenfors
William V. Gehrlein
Allan Gibbard
Edward J. Green
Paul Grout
Roger Guesnerie

Peter Hammond
Milton Harris
Richard Harris
Oliver Hart
Alan Haurie
Beth Hayes
Melvin Hinich
Gur Huberman
Tatsuro Ichiishi
Haruo Imai

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

Philip Jones
James Jordan
Kenneth Judd
Ehud Kalai
Mark Kamlet
Yoshi Kanemoto
David Kreps
Harold Kuhn
Finn Kydland

Guy Laroque
Lawrence J. Lau
David Livesey

Mark Machina
Michael Magill
Michael Maschler
Eric Maskin
Steven Matthews
Lionel W. McKenzie
Richard McLean
Paul Milgrom
James A. Mirrlees
Tapan Mitra
Hajime Miyazaki
James C. Moore
Dale T. Mortensen
Hervé Moulin
Roger Myerson

Abraham Neyman
William Novshek

Janusz Ordover
Martin Osborne

Edward Packel
Bezalel Peleg
Pierre Picard
H. M. Polemarchakis
Andrew Postlewaite
Edward Prescott
Daniel A. Primont

Arthur Raviv
Joe D. Reid
Jennifer F. Reinganum
Jeffrey Richelson
Donald K. Richter
Marcel K. Richter
John Riley
R. Tyrrell Rockafellar
Robert W. Rosenthal
James N. Rosse
Ariel Rubinstein
Harl E. Ryder

Mark A. Satterthwaite
Norman Schofield
Atle Seierstad
Larry Selden
Amartya Sen
Manimay Sengupta
Steven Shavell
L. A. Shepp
Robert J. Shiller
John B. Shoven
Steven Slutsky
Joel Sobel
Chester S. Spatt
Ross Starr
Nick Stern
Nancy Stokey
Philip Straffin
Chal Sussangkarn
Knut Sydsæter

Jacques Thisse
William Thomson
Robert M. Townsend

Neil Wallace
Henry Wan, Jr.
Laurence Weiss
Martin Weitzman
Louis Wilde
Charles A. Wilson
John D. Wilson

Menahem E. Yaari

Shmuel Zamir
Itzhak Zilcha

JOURNAL OF ECONOMIC THEORY 24, 155–156 (1981)

Papers to Appear in Forthcoming Issues

The following articles are among those newly accepted for publication in the *Journal of Economic Theory*. They are listed alphabetically by first author.

BETH ALLEN, Strict Rational Expectations Equilibria with Diffuseness. Correspondent: Professor Beth Allen, Department of Economics, University of Pennsylvania, 3718 Locust Walk CR, Philadelphia, PA 19104.

DAVID AUSTEN-SMITH, Restricted Pareto and Rights. Correspondent: Professor David Austen-Smith, Economics and Related Studies, University of York, Heslington, York YO1 5DD, England.

YVES BALASKO, Budget Constrained Pareto Efficient Allocations: A Dynamic Story. Correspondent: Professor Yves Balasko, CEPREMAP, 142, rue du Chevaleret, 75013 Paris, France.

STEVEN M. BARTA, Capital Accumulation in a Stochastic Decentralized Economy. Correspondent: Dr. Steven M. Barta, Room FJ 1B-124, Bell Laboratories, Holmdel, NJ 07733.

L. M. BENVENISTE AND J. A. SCHEINKMAN, Duality Theory for Dynamic Optimization Models of Economics: The Continuous Time Case. Correspondent: Professor J. A. Scheinkman, Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637.

JOHN CONLISK, The Law of Supply and Demand as a Law of Markov Chains. Correspondent: Professor John Conlisk, Department of Economics, University of California, San Diego, La Jolla, CA 92093.

MARTINE DUCHATELET, A Note on Increasing Returns to Scale and Learning by Doing. Correspondent: Professor Martine Duchatelet, Graduate School of Business, Columbia University, Uris Hall, #605B, New York, NY 10027.

SERGIU HART, The Number of Commodities Required to Represent a Market Game. Correspondent: Dr. Sergiu Hart, Department of Statistics, Tel-Aviv University, Tel-Aviv, Israel.

S. M. RAVI KANBUR, Increases in Risk with Kinked Payoff Functions. Correspondent: Professor S. M. Ravi Kanbur, Clare College, Cambridge CB2 1TL, England.

MORDECAI KURZ, Unemployment Equilibrium in an Economy with Linked Prices. Correspondent: Professor Mordecai Kurz, IMSSS, Stanford University, 4th Floor, Encina Hall, Stanford, CA 94305.

ALAIN LIPIETZ. The So-Called 'Transformation Problem' Revisited. Correspondent: Professor Alain Lipietz, CEPREMAP, 142, rue du Chevaleret, 75013 Paris, France.

MUKUL MAJUMDAR AND TAPAN MITRA, Intertemporal Allocation with a Non-Convex Technology: The Aggregative Framework. Correspondent: Professor Mukul Majumdar, Department of Economics, Cornell University, Ithaca, NY 14853.

LIONEL W. MCKENZIE, A Primal Route to the Turnpike and Liapounov Stability. Correspondent: Professor Lionel W. McKenzie, Department of Economics, University of Rochester, Rochester, NY 14627.

YOSHIHIKO OTANI AND JOSEPH SICILIAN, Equilibrium Allocations of Walrasian Preference Games. Correspondent: Professor Joseph Sicilian, Department of Economics, University of Kansas, 207 Summerfield, Lawrence, KS 660045.

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

Y. Y. Papageorgiou, A Probabilistic Model of Location Choice with Neighborhood Effects: Comment. Correspondent: Professor Y. Y. Papageorgiou, Department of Geography, McMaster University, 1280 Main Street, Hamilton, Ontario L8S 4K1 Canada.

Donald K. Richter, Weakly Democratic Regular Tax Equilibria in a Local Public Goods Economy with Perfect Consumer Mobility. Correspondent: Professor Donald K. Richter, Economics, Boston College, Chestnut Hill, MA 02167.

Neil Wallace, A Hybrid Fiat-Commodity Monetary System. Correspondent: Dr. Neil Wallace, Federal Reverse Bank, 250 Marquette Avenue, Minneapolis, MN 55480.

Claus Weddepohl, Equilibria with Rationing in an Economy with Increasing Returns. Correspondent: Professor Claus Weddepohl, Faculty of Actuarial Science and Econometrics, Jodenbreestraat 23, 1011 NH Amsterdam.

Printed in Belgium

JOURNAL OF ECONOMIC THEORY 24, 157–167 (1981)

On Uniqueness of a Steady State and Convergence of Optimal Paths in Multisector Models of Optimal Growth with a Discount Rate

KIYOHIKO GIICHI NISHIMURA

Department of Economics, Yale University, New Haven, Connecticut 06520

Received September 11, 1978; revised April 15, 1980

1. INTRODUCTION

It is well known that in multisector models of optimal growth, optimal paths converge to a unique steady state when future utilities are not discounted and utility functions are strictly concave. See Gale [7], McKenzie [13], and Brock [2] for discrete time models, and Rockafellar [15] for continuous time models. Scheinkman [18] shows uniqueness of a steady state and convergence of optimal paths in discounted optimal growth models when utility functions are separable on state variables and control variables.

For non-separable utility functions, there are several examples of multiple steady states and unstable optimal paths. Kurz [8] and Liviatan and Samuelson [11] obtain multiple steady states. Kurz [9] suggests the possibility of instability, while Sutherland [17] produces an example of unstable optimal paths. These examples show that the strict concavity assumption on utility functions is not sufficient to obtain either uniqueness or convergence in non-separable utility models.

In the present state of knowledge, global curvature conditions on derived Hamiltonian functions (see Cass and Shell [6], Rockafellar [16], and Brock and Scheinkman [4]) are the only sufficient conditions for convergence and uniqueness available in general non-separable utility models, though Brock and Scheinkman [5] produce several useful conditions in special cases. Lau [10] establishes formulae relating the Hessian matrix of utility functions to that of derived Hamiltonian functions, so that the global curvature conditions, such as the Q matrix condition of Brock and Scheinkman [4], can be expressed in terms of conditions on utility functions. Derived conditions on utility functions, however, are complicated and hard to interpret if there are more than two state variables. This complexity might

0022-0531/81/020157–11$02.00/0

Copyright © 1981

prevent economists from utilizing the global curvature conditions when they face optimal growth problems with non-separable utility functions.

The purpose of this paper is to obtain several sets of sufficient conditions for uniqueness and convergence based on utility functions, not on the derived Hamiltonian functions. Although we utilize the value-loss technique which Cass and Shell and others use in their works, we do not adopt their Hamiltonian formulation. Instead, we develop a new method, which enables us to investigate directly effects of the shape of utility functions on uniqueness of a steady state and convergence of optimal paths. This method is based on the classical calculus of variations and employs concavity property of utility functions.

Section 2 presents the model and theorems, and discusses assumptions. Proofs of the theorems are given in Section 3. Section 4 contains an illustrative example which compares the results in this paper with those in other previous works. Some concluding remarks are given in Section 5.

2. MAIN RESULTS

We investigate the following optimal growth problems with a positive discount rate:

$$\text{Maximize} \int_0^{\infty} U(\dot{k}, k)\, e^{-\rho t}\, dt \qquad \text{s.t.} \qquad k(0) = k_0. \tag{P}$$

Here $U: B \subset R^n \times R^n \to R$ is C^2, concave, non-separable on $\dot{k}$ and k, and B is compact and convex with non-empty interior. Problem (P) allows the existence of joint production, wealth effect, and adjustment costs. Concavity of U is generated by convexity of technology and preference. Rockafellar [15] shows that if $\rho = 0$, optimal paths converge to a unique steady state of the Euler differential equations, which are necessary conditions for optimality. Scheinkman [18] proves convergence of optimal paths when U is separable on $\dot{k}$ and k. This paper investigates sufficient conditions for convergence and uniqueness with non-separable utility functions.

In order to concentrate our analysis on the behavior of optimal paths, we simply assume the existence of an interior optimal path from k_0 and an interior optimal steady state in the problem (P). B^0 denotes the interior of B.

ASSUMPTION (A.1). *There exists an interior optimal path $k^* : (\dot{k}^*, k^*) \in B^0$ such that*

$$W(k_0) \equiv \int_0^{\infty} U(\dot{k}^*, k^*) e^{-\rho t}\, dt < +\infty,$$

where $k^*(0) = k_0$.

Assumption (A.2). *There exists an optimal steady state* $\bar{k} : (0, \bar{k}) \in B^0$ *such that if* $k_0 = \bar{k}$, *then* $k^* = \bar{k}$ *for all* $t \geqslant 0$.

From (A.1), k^* *must satisfy the following Euler differential equations which are necessary conditions for optimality. Hereafter, for a function* $F(x, y)$, F_1 *and* F_2 *will denote* $\partial F/\partial x$ *and* $\partial F/\partial y$, *respectively. Similarly,* $F_{11} \equiv \partial^2 F/\partial x^2$ *and* $F_{12} \equiv \partial^2 F/\partial x \partial y$.

Euler Differential Equations

$$U_1(\dot{k}^*, k^*) = q^* \text{ (Definition of } q^*\text{)},$$
$$U_2(\dot{k}^*, k^*) + \rho q^* = \dot{q}^*. \tag{E}$$

Similarly, $\bar{k}$ must satisfy

$$U_1(0, \bar{k}) = \bar{q} \text{ (Definition of } \bar{q}\text{)},$$
$$U_2(0, \bar{k}) + \rho \bar{q} = 0. \tag{ES}$$

Before proceeding with our analysis we introduce a partial order on the concavity of functions which will be central in our later analysis.

Definition. Let f and $g: D \subset R^n \times R^n \to R$ be C^1 and concave, where D is convex with non-empty interior. $f(x, y)$ is *at least as concave as* $g(x, y)$ on D if $f(x, y) - g(x, y)$ is a concave function on D.

The following property can easily be proved. $|\ |$ denotes the Euclidean norm and T is the transpose.

$U(x, y)$ is at least as concave as

$$-\tfrac{1}{2}\{\alpha x^T x + \beta y^T y + \eta x^T y\}, \qquad \text{where } 4\alpha\beta > \eta^2$$

$$\Leftrightarrow$$

$$\begin{aligned} &U(x', y') - U(x, y) \\ &\quad \leqslant U_1(x, y)\{x' - x\} + U_2(x, y)\{y' - y\} \\ &\qquad - \tfrac{1}{2}\{\alpha |x' - x|^2 + \beta |y' - y|^2 + \eta (x' - x)^T (y' - y)\}. \end{aligned}$$

It should be noted here that this definition is an extention of Rockafellar's α-concavity in [16]. More precisely, α-concavity of $f(x, y)$ is equivalent to $f(x, y)$ being at least as concave as $-\tfrac{1}{2}\alpha\{x^T x + y^T y\}$. If $U(x, y)$ is at least as concave as $\tfrac{1}{2}\{\alpha x^T x + \beta y^T y + \eta x^T y\}$, where $4\alpha\beta > \eta^2$, then $U(x, y)$ is strictly concave.

The following new assumptions on U are of utmost importance throughout this paper. In these assumptions, (x, y) and (x', y') are arbitrary points in B^0.

Assumption (A.3). *There exist a number* η *and positive numbers* α, β, γ, *and* δ *such that*

(N.1) $4\alpha\beta > \eta^2$,

(N.2) $4\alpha(\beta - \delta) > \gamma^2$,

(N.3) $U(x, y)$ *is at least as concave as* $-\frac{1}{2}\{\alpha x^T x + \beta y^T y + \eta x^T y\}$,

(N.4) *Let* $F(x, y) \equiv \rho U(x, y) - \frac{1}{2}\eta x^T x$. *Then*

$$|F_1(x', y') - F_1(x, y)| \leqslant \gamma |x' - x| + \delta |y' - y|.$$

If U satisfies (N.1) and (N.3), then U is strictly concave. α and β can be regarded as lower bounds on the degree of concavity of U. Part (N.4) places upper bounds on the elements of U_{11} and U_{12}, these bounds being dependent on the discount rate and η. Because U is concave, the upper bound on the elements of U_{11} involves an upper bound on the degree of concavity of U. Thus (N.2) requires that U should be sufficiently concave (α and β must be large enough), but not excessively concave (γ must not be excessively large) compared with the discount rate ρ.

In some cases, we may not be able to find η, α, β, γ, and δ satisfying (A.3). Then the following (A.4) may be useful.

Assumption (A.4). (i) *There exist a number* η *and positive numbers* α, β, γ, *and* δ *satisfying* (N.1), (N.3), (N.4), *and* (N.2′) $4\alpha(\beta - \delta) \geqslant \gamma^2$.

(ii) $U(x, y)$ *is* C^3 *and* U_{11} *is non-singular on* B^0.

(iii) *Except for* $\bar{k}$ *defined in* (A.2), *no optimal path satisfies the following differential equations.*

$$\{U_{11}(\dot{k}, k)\ddot{k} + U_{12}(\dot{k}, k)\dot{k}\}\{\bar{k} - k\} = \{U_1(\dot{k}, k) - U_1(0, \bar{k})\}\dot{k}.$$

Assumption (A.4) allows equality in (N.2′), though it needs extra assumptions (ii) and (iii). It is not difficult to tell whether U satisfies (ii) or not. Although (iii) looks like a complicated condition, it is easy in some models to prove that U satisfies (iii) by using the phase diagram method. One example of the use of (A.3) and (A.4) will be given in Section 4.

In the following section, we prove two theorems.

Theorem 1. *Assume* (A.1), (A.2), *and* (A.3). *Then an optimal steady state* $\bar{k}$ *is unique and optimal paths* k^* *converge to* $\bar{k}$.

Theorem 2. *Assume* (A.1), (A.2), *and* (A.4). *Then an optimal steady state* $\bar{k}$ *is unique and optimal paths* k^* *converge to* $\bar{k}$.

3. PROOFS OF THEOREMS

3.1. *Theorem 1*

The first point to notice is that (A.3) implies strict concavity of U. Thus an optimal path from k_0 is unique.

DEFINITION 1. Let k and k' be optimal paths from k_0 and k'_0, respectively, where $k_0 \neq k'_0$.

$$\begin{aligned} q &\equiv U_1(\dot{k}, k), \\ q' &\equiv U_1(\dot{k}', k'), \\ V &\equiv \{q - q'\}\{k' - k\}. \end{aligned}$$

LEMMA 1. *$V \leqslant 0$ for all $t \geqslant 0$.*

Proof. From (A.1) and concavity of U, $W(k)$ is concave, differentiable, and $W'(k) = -q$ for $t \geqslant 0$, where $W' \equiv \partial W/\partial k$. (See Benveniste and Scheinkman [1] for details.) Because of the concavity of W, we obtain

$$\{W'(k') - W'(k)\}\{k' - k\} \leqslant 0,$$

which implies

$$\{-q' + q\}\{k' - k\} = \{q - q'\}\{k' - k\} = V \leqslant 0. \qquad \text{Q.E.D.}$$

LEMMA 2.

$$\dot{V} \geqslant \{\eta(\dot{k}' - \dot{k})^T + \rho(q - q')\}\{k' - k\} + \alpha\,|\dot{k}' - \dot{k}|^2 + \beta\,|k' - k|^2. \tag{1}$$

Proof. From (N.3), we have the following relations.

$$\begin{aligned} &U(\dot{k}', k') - U(\dot{k}, k) \\ &\quad \leqslant U_1(\dot{k}, k)\{\dot{k}' - \dot{k}\} + U_2(\dot{k}, k)\{k' - k\} \\ &\qquad - \tfrac{1}{2}\{\alpha\,|\dot{k}' - \dot{k}|^2 + \beta\,|k' - k|^2 + \eta(\dot{k}' - \dot{k})^T(k' - k)\}, \end{aligned} \tag{2}$$

$$\begin{aligned} &U(\dot{k}, k) - U(\dot{k}', k') \\ &\quad \leqslant U_1(\dot{k}', k')\{\dot{k} - \dot{k}'\} + U_2(\dot{k}', k')\{k - k'\} \\ &\qquad - \tfrac{1}{2}\{\alpha\,|\dot{k} - \dot{k}'|^2 + \beta\,|k - k'|^2 + \eta(\dot{k} - \dot{k}')^T(k - k')\}. \end{aligned} \tag{3}$$

Adding up both sides of (2) and (3), and using (E) and Definition 1, we obtain

$$\{q-q'\}\{\dot{k}'-\dot{k}\}+\{\dot{q}-\dot{q}'\}\{k'-k\}$$
$$\geqslant \{\eta(\dot{k}'-\dot{k})^T+\rho(q-q')\}\{k'-k\}+\alpha\,|\dot{k}'-\dot{k}|^2+\beta\,|k'-k|^2. \quad \text{Q.E.D.}$$

LEMMA 3.

$$\dot{V}\geqslant \sigma\{|\dot{k}'-\dot{k}|+|k'-k|\}^2, \tag{4}$$

where $\sigma=\{4\alpha(\beta-\delta)-\gamma^2\}/\{\alpha+\beta-\delta+\gamma\}>0.$

Proof. From (N.4) we obtain

$$\begin{aligned}
&|\eta(\dot{k}'-\dot{k})^T+\rho(q-q')|\\
&\quad=|\{\rho U_1(\dot{k}',k')-\eta\dot{k}'^T\}-\{\rho U_1(\dot{k},k)-\eta\dot{k}^T\}|\\
&\quad=|F_1(\dot{k}',k')-F_1(\dot{k},k)|\\
&\quad\leqslant\gamma\,|\dot{k}'-\dot{k}|+\delta\,|k'-k|.
\end{aligned}\tag{5}$$

By using (1), (5), and the fact that $z\geqslant -|z|$ for any number z, we obtain the following transformations easily.

$$\begin{aligned}
\dot{V}&\geqslant -|\eta(\dot{k}'-\dot{k})^T+\rho(q-q')|\,|k'-k|+\alpha\,|\dot{k}'-\dot{k}|^2+\beta\,|k'-k|^2\\
&\geqslant -\{\gamma\,|\dot{k}'-\dot{k}|+\delta\,|k'-k|\}\,|k'-k|+\alpha\,|\dot{k}'-\dot{k}|^2+\beta\,|k'-k|^2\\
&=\alpha\,|\dot{k}'-\dot{k}|^2-\gamma\,|\dot{k}'-\dot{k}|\,|k'-k|+(\beta-\delta)\,|k'-k|^2.
\end{aligned}\tag{6}$$

The following inequality is well known: For all real numbers $\lambda\geqslant 0$ and $\mu\geqslant 0$,

$$a\lambda^2-c\lambda\mu+b\mu^2\geqslant\{(4ab-c^2)/4(a+b+c)\}(\lambda+\mu)^2, \tag{7}$$

if $a>0$, $b>0$, $c>0$, and $4ab\geqslant c^2$. (See Rockafellar [16, p. 87].) By using this inequality and (N.2), we obtain (4) from (6). Q.E.D.

Proof of Theorem 1. (i) Suppose that there be two different optimal steady states $\bar{k}$ and $\bar{k}'$. From (6) we obtain

$$0=\dot{V}\geqslant(\beta-\delta)\,|\bar{k}'-\bar{k}|^2>0, \tag{8}$$

where we use the property that $\bar{k}'$ and $\bar{k}$ are constant over time. It is evident that (8) is a contradiction, so that $\bar{k}$ must be unique.

(ii) Let $V^*\equiv\{q-\bar{q}\}\{\bar{k}-k\}$, where $\bar{q}\equiv U_1(0,\bar{k})$. From Lemma 3, if $|\bar{k}-k|>\varepsilon$, then $\dot{V}^*>\sigma\varepsilon$, where $\sigma>0$. By using the same routine argument as in Cass and Shell [6, p. 59], we can show that if k does not converge to $\bar{k}$, this contradicts Lemma 1. Q.E.D.

3.2. *Theorem 2*

Equation (6) is still true under (A.4). Consequently, the proof of uniqueness of $\bar{k}$ under (A.4) is exactly the same as under (A.3).

As before, (A.4) implies strict concavity of U, so that optimal path k^* from k_0 is unique.

In order to prove the convergence property of k^*, we use the following theorem due to La Salle. (See Lefschetz [19, Chap. 9].)

La Salle's Theorem. *Let $x(t)$ be a bounded solution of $\dot{x} = G(x)$ defined on $t \in [0, +\infty)$, where $G: R^n \to R^n$ is C^1. If there exists a C^1 function $L(x): R^n \to R$ such that, along with $x(t)$,*

(a) *For some $\bar{x}$, $L(\bar{x}) = 0$, and if $x(t) \neq \bar{x}$, then $L(x(t)) > 0$,*

(b) *$\dot{L}(x(t)) \leqslant 0$, and*

(c) *Except for $x(t) = \bar{x}$ for $t \geqslant 0$, $x(t)$ does not satisfy $\dot{L}(x(t)) = 0$,*

then $x(t)$ converges $\bar{x}$.

We cannot apply La Salle's theorem directly to (E), because (E) are not first order differential equations. However, (E) can be rewritten as

$$U_2(\dot{k}^*, k^*) + \rho U_1(\dot{k}^*, k^*) = U_{11}(\dot{k}^*, k^*)\ddot{k}^* + U_{12}(\dot{k}^*, k^*)\dot{k}^*, \tag{9}$$

which can be transformed into

$$\begin{aligned} \ddot{k}^* &= U_{11}^{-1}(\dot{k}^*, k^*)\{U_2(\dot{k}^*, k^*) + \rho U_1(\dot{k}^*, k^*) - U_{12}(\dot{k}^*, k^*)\dot{k}^*\} \\ &\equiv f(\dot{k}^*, k^*). \end{aligned} \tag{10}$$

$: B^0 \to R^n$ is C^1 because of (ii) of (A.4).

Definition 2. $z \equiv \dot{k}^*$, $k \equiv k^*$, and

$$v(z, k) \equiv \{U_1(z, k) - U_1(0, \bar{k})\}\{\bar{k} - k\}.$$

By using Definition 2, we obtain from (10)

$$\dot{z} = f(z, k), \tag{11}$$

$$\dot{k} = z. \tag{12}$$

Thus we get autonomous first order differential equations with the following initial conditions:

$$k(0) = k^*(0, k_0) = k_0, \tag{13}$$

$$z(0) = \dot{k}^*(0, k_0). \tag{14}$$

In the following, we show that $-V(z, k)$ is a Liapunov function in La Salle's theorem.

LEMMA 4. *$V(z, k) = 0$ iff $(z, k) = (0, \bar{k})$, and $V(z, k) < 0$ otherwise.*

Proof. By using the same argument as the one in Lemma 1, we can show that $V(z, k)$ is non-positive. U is strictly concave and B^0 is convex, so that W is strictly concave. Thus we obtain for $k^* \neq \bar{k}$,

$$\begin{aligned} W(k^*) - W(\bar{k}) &< W'(\bar{k})\{k^* - \bar{k}\}, \\ W(\bar{k}) - W(k^*) &< W'(k^*)\{\bar{k} - k^*\}, \end{aligned} \tag{15}$$

which yields

$$\begin{aligned} 0 &< \{W'(\bar{k}) - W'(k^*)\}\{k^* - \bar{k}\} \\ &= \{U_1(0, \bar{k}) - U_1(\dot{k}^*, k^*)\}\{\bar{k} - k^*\} = -V(z, k). \end{aligned}$$

Here we use use Definition 2. Thus if $k \neq \bar{k}$, then $V(z, k) < 0$. By definition, if $(z, k) = (0, \bar{k})$, then $V(z, k) = 0$. This ends the proof. Q.E.D.

LEMMA 5. *$\dot{V}(z, k) \geqslant 0$.*

Proof. Equation (6) is still true under (A.4). It is easy to prove that (i) of (A.4) is sufficient to have $\dot{V}(z, k) \geqslant 0$. Q.E.D.

Proof of Theorem 2. The proof of uniqueness of $\bar{k}$ is the same as the one in Theorem 1. Now we show that all conditions of La Salle's theorem are satisfied.

First, the differential equations (11), (12), (13), and (14) are first order differential equations, and $f(z, k)$ and z are C^1 with respect to z and k. From (A.1) and Definition 2, (z, k) is bounded and defined on $t \in [0, +\infty)$.

Second, V is C^1 with respect to z and k because of (ii) of (A.4). Lemmas 4 and 5 show that $-V$ satisfies (a) and (b) of La Salle's theorem. Part (iii) of (A.4) implies (c).

Thus all the conditions of La Salle's theorem are satisfied, so that k converges to $\bar{k}$, which is equivalent to convergence of k^* to $\bar{k}$. Q.E.D.

4. An Example

In this section, we present an example which illustrates the relationship between (A.3), (A.4), and other sufficient conditions derived by Cass and Shell and others.

Consider the one-dimensional quadratic utility function

$$U(\dot{k}, k) \equiv a\dot{k}k - \tfrac{1}{2}b\dot{k}^2 - \tfrac{1}{2}ck^2$$

defined on the set

$$|\dot{k}| \leqslant A \qquad \text{and} \qquad |k| \leqslant B,$$

where $b > 0$, $c > 0$, and $bc > a^2$, and A and B are some positive real numbers.

Because there is only one state variable, the global curvature conditions of Cass and Shell, Brock and Scheinkman, and Rockafellar can easily be applied in this simple case. These conditions reduce to the inequality

$$4(bc - a^2) > \rho^2 b^2. \tag{17}$$

First, we show that the global curvature conditions (17) are sufficient conditions for (A.3) in this simple example. Take $\alpha = b$, $\beta = c$, and $\eta = -2a$. Then it is easy to prove that if U satisfies (17), it does (A.3).

Next, we prove that (A.3) is a sufficient condition for (A.4) in this example. It is evident that U is C^3 and $U_{11} \neq 0$. Thus (ii) of (A.4) holds. By using phase diagrams, it is fairly easy to prove that (iii) of (A.4) is true. Parts (N.1) through (N.4) are sufficient conditions for (i) of (A.4). Consequently, (A.3) is a sufficient condition for (A.4) in this case.

Finally, we present an example in which (A.4) holds, while (A.3) and (17) do not. Consider the following utility function and the discount rate:

$$U(\dot{k}, k) \equiv \dot{k}k - \dot{k}^2 - k^2 \qquad \text{and} \qquad \rho = 3^{1/2}.$$

It is easy to show that both (17) and (A.3) fail, but (A.4) holds.

5. Concluding Remarks

In this paper we have developed a new method for investigating the uniqueness of a steady state and convergence of optimal paths in multisector models of optimal growth having a non-separable utility function and a positive discount rate. By means of this method we obtained two sets of sufficient conditions for uniqueness and convergence on curvature of utility functions. These conditions are shown to be superior (A.4), or at least not inferior (A.3) to the global curvature conditions on derived Hamiltonian functions if utility functions are one-dimensional and quadratic.

For more general cases, however, we have been unable to establish the relationship between the new conditions and the global curvature conditions on Hamiltonian functions, though Nishimura [21] suggests a close relationship between them. The investigation of the relationship and the search for new stability conditions seem an important task for future research.

With our analysis we obtained sufficient conditions for convergence to a point, asymptotic stability. In economics, however, asymptotic stability is merely one of many meaningful stability concepts. For example, convergence to a limit cycle may be more interesting than simple asymptotic stability if one would like to apply optimal growth theory to the theory of business cycles. Thus a more general notion of long-run behavior of an economy needs to be formulated and to be applied in optimal growth models. Preliminary results of research in this direction are reported in Nishimura [20].

ACKNOWLEDGMENTS

I am indebted to W. A. Brock, J. Benhabib, Kazuo Nishimura, and K. Asako for their helpful comments. I am especially grateful to K. Iwai and the anonymous referee for their valuable suggestions. Any errors that may remain, however, are entirely my own. Financial support from Murata Overseas Scholarship Foundation of Japan is gratefully acknowledged.

REFERENCES

1. L. M. Benveniste and J. A. Scheinkman, On the differentiability of the value function in dynamic models of economics, *Econometrica* **47** (1979), 727–732.
2. W. A. Brock, On existence of weakly maximal programmes in a multisector economy, *Rev. Econ. Studies* **37** (1970), 275–280.
3. W. A. Brock, The global asymptotic stability of optimal control: A survey of recent results, *in* "Frontiers of Quantitative Economics, Vol. IIIB" (M. Intriligator, Ed.), pp. 207–237, North-Holland, Amsterdam, 1977.
4. W. A. Brock and J. A. Scheinkman, Global asymptotic stability of optimal control systems with applications to the theory of economic growth, *J. Econ. Theory* **12** (1976), 164–190.
5. W. A. Brock and J. A. Scheinkman, The global asymptotic stability of optimal control with applications to dynamic economic theory, *in* "Applications of Control Theory to Economic Analysis" (J. Pitchford and S. Turnovsky, Eds.), North-Holland, Amsterdam, 1977.
6. D. Cass and K. Shell, The structure and stability of competitive dynamical systems, *J. Econ. Theory* **12** (1976), 31–70.
7. D. Gale, On optimal development in a multisector economy, *Rev. Econ. Studies* **34** (1967), 1–18.

8. M. KURZ, Optimal economic growth and wealth effects, *Internat. Econ. Rev.* **9** (1968), 348–357.
9. M. KURZ, The general instability of a class of competitive growth processes, *Rev. Econ. Studies* **35** (1968), 155–174.
10. L. LAU, A characterization of the normalized restricted profit function, *J. Econ. Theory* **12** (1976), 131–163.
11. N. LIVIATAN AND P. A. SAMUELSON, Notes on turnpikes: Stable and unstable, *J. Econ. Theory* **1** (1969), 454–475.
12. M. J. P. MAGILL, Some new results on the local stability of the process of capital accumulation, *J. Econ. Theory* **15** (1977), 174–210.
13. L. W. MCKENZIE, Accumulation programs of maximal utility and the von Neumann facet, *in* "Value, Capital and Growth, Papers in Honor of Sir John Hicks" (J. N. Wolfe, Ed.), pp. 353–383, Edinburgh Univ. Press, Edinburgh, 1968.
14. L. W. MCKENZIE, Turnpike theory, *Econometrica* **44** (1976), 841–865.
15. R. T. ROCKAFELLAR, Saddle points of Hamiltonian systems in convex problems of Lagrange, *J. Optimization Theory Appl.* **12** (1973), 367–390.
16. R. T. ROCKAFELLAR, Saddle points of Hamiltonian systems in convex Lagrange problems having a nonzero discount rate, *J. Econ. Theory* **12** (1976), 71–113.
17. W. R. S. SUTHERLAND, On optimal development in a multisectoral economy: The discounted case, *Rev. Econ. Studies* **37** (1970), 585–589.
18. J. A. SCHEINKMAN, Stability of separable Hamiltonians and investment theory, *Rev. Econ. Studies* **45** (1978), 559–570.
19. S. LEFSCHETZ, Stability of Nonlinear Control Systems, Academic Press, New York, 1965.
20. K. G. NISHIMURA, A new concept of stability and the theory of optimal growth, mimeo, Yale University, March 1979.
21. K. G. NISHIMURA, A note on the global stability in dynamical economic systems, *Econ. Lett.* **4** (1979), 95–98.

JOURNAL OF ECONOMIC THEORY 24, 168–186 (1981)

Optimal Growth and Uncertainty: The Borrowing Models*

L. C. MACLEAN

Faculty of Business Administration, Dalhousie University,
Halifax, Nova Scotia B3H 4H8, Canada

AND

C. A. FIELD AND W. R. S. SUTHERLAND

Mathematics Department, Dalhousie University,
Halifax, Nova Scotia B3H 4H8, Canada

Received September 21, 1978; revised September 4, 1979

1. INTRODUCTION AND SUMMARY

The theory of optimal economic growth is concerned with the problem of how to best allocate some current stock of capital between capital for immediate consumption and capital for future investment. This allocation is normally required to satisfy a strong type of budget constraint at each point in time; deficits are not allowed and surpluses cannot be retained for future use.

In this paper we examine the same allocation question under a weaker type of constraint. Deficits and surpluses will now be carried forward in time, but only the total deficit over the planning horizon will be constrained. A decision to incur a deficit (surplus) will be called a borrowing (lending) activity.

A similar concept of borrowing appears in the planning models which arise in the theory of the consumer. There an individual must decide between consumption and investment at each point of time in his planning horizon or lifetime. Borrowing, as well as being an economic acitivity, also serves as a device for introducing a concept of life insurance. Yaari [10] discusses the qualitative differences between consumer models with and without insurance. Hakansson [14] gives explicit optimal policies for a specific class of utility functions.

---

* This research was supported in part by grants from the National Research Council of Canada.

0022-0531/81/020168-19$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved.

Returning to the economic growth literature, our starting point is the one-sector nonlinear deterministic production model of Koopmans [5], Cass [2] and Gale [3]. The existence of optimal investment and consumption policies and their qualitative properties are well known. The turnpike (or golden-rule) phenomena of long-run investment being the most celebrated results. Brock and Mirman [1] extended this theory to the case of a stochastic production process and showed that there exists a stationary distribution (over the set of turnpike states) which is approached by long-run optimal investment policies.

In Section 2, the borrowing activity is introduced in a discrete-time finite-horizon economic growth model with deterministic production process. The production process is then extended to the case of random or uncertain outputs. The activity of intertemporal borrowing (or lending) of capital involves a fixed rate of interest but is otherwise unrestricted and costless. Capital is borrowed from (or lent to) some unspecified exogeneous source. Two versions of the model are considered. If only the average (or expected) net borrowing over the planning horizon must be repaid, we have the *optimistic* model (of Section 3). However, if the actual net borrowing must be repaid, we have the *pessimistic* model (of Section 4). Roughly speaking, these two versions of the borrowing model correspond to versions of Yaari's consumer model depending on whether life insurance is available or not. However, unlike these consumer models, we do not postulate a specific mechanism for insurance. Our exogeneous source is unspecified. It faces no risks in a pessimistic model and acts as the holder of any final deficits or surpluses in an optimistic model. The latter will be in balance if it were to serve as an exchange bank for a large number of identical economies.

An alternative justification for using such an exogeneous source can be based on the idea of "rolling plans." Here an economy revises its policies after each time period while advancing its planning horizon ahead by one period. Using an optimistic policy, the economy always expects to repay its debts, the exogeneous source always expects to have a zero net balance, and the day of final reckoning is ever deferred.

Sections 3 and 4 describe the optimal investment and consumption policies for the optimistic and pessimistic models. Not surprizingly, the optimistic model is the simpler version. Its decisions on investment versus consumption are deterministic with borrowing absorbing all the stochastic behaviour. Section 5 presents explicit formulas for the consumption policies in the two versions for the special case of a quadratic utility function.

2. The Deterministic Borrowing Model

Gale [3] considers a discrete-time one-sector economic growth model. There is a single good (capital) which can either be invested or consumed. An investment x_t in time period t produces an output $f(x_t)$ in period $t+1$. Consumption c_t yields utility $u(c_t)$ during period t. The initial capital stock is y_0, the planning horizon T periods, the final stock is y_T. An optimal policy is a sequence (x_t, c_t) of (non-negative) investments and consumptions which satisfy the budget constraints and yield a maximum for the sum of the discounted utilities.

We now introduce a third activity, z_t, which represents borrowed capital (if negative) or repayment of capital (if positive) with respect to some unspecified exogeneous stock. Interest on such activities is represented by a factor $\rho > 0$; discounting by $\delta > 0$. This modification yields the following *borrowing model* (which reduces to Gale's model if all $z_t = 0$):

Find $x_t \geqslant 0$, $c_t \geqslant 0$ and z_t which maximize

$\sum_{t=1}^{T} \delta^{t-1} u(c_t)$ subject to the budget constraints

$$\begin{aligned} &x_1 + c_1 + z_1 \leqslant y_0, \\ &x_t + c_t + z_t \leqslant f(x_{t-1}), \qquad t = 2, \ldots, T, \\ &y_T + z_{T+1} \leqslant f(x_T), \end{aligned}$$

and the repayment constraint

$$\sum_{t=1}^{T+1} \rho^{t-1} z_t \geqslant 0.$$

The properties of this and related deterministic borrowing models are given elsewhere (see [8]). Our present purpose is to consider the case of a stochastic production process. The single production function f is now replaced by a family f_ω, where $\omega \in \Omega$ and $(\Omega, \mathscr{B}, \mathscr{P})$ is a probability space. Thus the output $Y(\omega) = F(x) = f_\omega(x)$ is a random variable, for each fixed input x. If the input is a random variable X, then $Y = F(X)$ is a random variable on the probability space $(\Omega \times \Omega,\ \mathscr{B} \times \mathscr{B},\ \mathscr{P} \times \mathscr{P})$ with $Y(\omega_1, \omega_2) = f_{\omega_2}(X(\omega_1))$.

The stochastic production process is assumed to be stationary in that the same probability space is used in all time periods. We also assume that, for each $\omega \in \Omega$, $f_\omega(x)$ is an increasing concave function such that $f_\omega(0) = 0$ and $\rho f_\omega(x) - x \leqslant M$ holds for all $x \geqslant 0$ for some bound M. The utility function $u(c)$ is also assumed to be an increasing, strictly concave function for $c > 0$. For notational convenience the same $\omega_{T+1} \in \Omega$ is used for both the final period of the production process and the final stocks Y_T. As usual $E(W)$ represents the expected value of a random variable W.

3. The Optimistic Borrowing Model

When the production process is stochastic the activities x_t, c_t, z_t become random variables with respect to the product probability space $(\Omega^t, \mathscr{B}^t, \mathscr{P}^t)$. Letting $\omega^t = (\omega_1, ..., \omega_t)$, these variables are denoted by $X_t(\omega_1, ..., \omega_t) = X_t(\omega^t)$, $C_t(\omega_1, ..., \omega_t) = C_t(\omega^t)$, $Z_t(\omega_1, ..., \omega_t) = Z_t(\omega^t)$ or simply as X_t, C_t, Z_t. The production process is denoted by $F(X) = f_\omega(X)$.

The *optimistic borrowing model* is defined by the following problem:

> Find random variables $X_t(\omega^t) \geqslant 0$, $C_t(\omega^t) \geqslant 0$ and $Z_t(\omega^t)$ which maximize $\sum_{t=1}^{T} \delta^{t-1} E(u(C_t))$ subject to the budget constraints
>
> $$X_1 + C_1 + Z_1 \leqslant Y_0,$$
> $$X_t + C_t + Z_t \leqslant F(X_{t-1}) \qquad \text{for} \quad t = 2, ..., T,$$
> $$Y_T + Z_{T+1} \leqslant F(X_T),$$
>
> and the repayment constraint
>
> $$\sum_{t=1}^{T+1} \rho^{t-1} E(Z_t) \geqslant 0.$$

All constraints are required to hold pointwise. That is, the typical budget constraint is

$$X_t(\omega^t) + C_t(\omega^t) + Z_t(\omega^t) \leqslant f_{\omega_t}(X_{t-1}(\omega^{t-1}))$$

for all $\omega^t \in \Omega^t$. Note that the initial and final stocks are considered to be random variables, $Y_0(\omega^1)$ and $Y_T(\omega^{T+1})$. A sequence (X_t, C_t, Z_t) which satisfies all of the constraints (for all $\omega^{T+1} \in \Omega^{T+1}$) is said to be a *feasible T-period policy* from Y_0 to Y_T.

Since $Z_t(\omega_t) > 0$ is interpreted as a repayment ($Z_t(\omega_t) < 0$ being borrowing), then the repayment constraint requires only that the *expected* total borrowing be repaid. The pessimistic model (discussed in Section 4) differs from this model only in requiring that the actual total borrowing be repaid. The effect of introducing a borrowing activity in a stochastic economic growth model is to relax the budget constraints and hence yield optimal policies which are more or less insensitive to the randomness inherent in the production process.

Let $f(x) = E(F(x))$ denote the expected (or average) production function. Our first result shows that the optimal investment and consumption policies are deterministic in the optimistic model. Moreover, they depend only on the average production function f, the expected initial and final stocks $E(Y_0)$, $E(Y_T)$ and the factors δ and ρ.

THEOREM 1. *If (x_t, c_t, z_t) is an optimal T-period policy from $y_0 = E(Y_0)$ to $y_T = E(Y_T)$ in the deterministic borrowing model with production function $f = E(F)$, then (x_t, c_t, Z_t) is an optimal T-period policy from Y_0 to Y_T for the optimistic model, where*

$$\begin{aligned} Z_1(\omega^1) &= Y_0(\omega^1) - x_1 - c_1, \\ Z_t(\omega^t) &= f_{\omega_t}(x_{t-1}) - x_t - c_t \qquad \text{for} \quad t = 2, \ldots, T, \\ Z_{T+1}(\omega^{T+1}) &= f_{\omega_{T+1}}(x_T) - Y_T(\omega^{T+1}). \end{aligned}$$

Proof. Let (x_t, c_t, z_t) be any feasible policy for the deterministic model and Z_t as defined, then it is easily verified that (x_t, c_t, Z_t) is feasible for the optimistic model. Conversely, if (X_t, C_t, Z_t) is feasible for the optimistic model from Y_0 to Y_T then $x_t = E(X_t)$, $c_t = E(C_t)$, $z_t = E(Z_t)$ is feasible for the deterministic model from y_0 to y_T. This follows from the concavity of F and Jensen's inequality as $E(X_t) + E(C_t) + E(Z_t) \leqslant E(F(X_{t-1})) \leqslant f(E(X_{t-1}))$; also $E(Y_T) + E(Z_{T+1}) \leqslant E(F(X_T)) \leqslant f(E(X_T))$. Now let (X_t, C_t, Z_t) be any feasible policy from Y_0 to Y_T, then applying Jensen's inequality to the concave function u, we obtain

$$\sum_{t=1}^{T} \delta^{t-1} E(u(C_t)) \leqslant \sum_{t=1}^{T} \delta^{t-1} u(E(C_t)) \leqslant \sum_{t=1}^{T} \delta^{t-1} u(c_t)$$

since (x_t, c_t, z_t) is the optimal deterministic policy, which completes the proof. ∎

Since the optimistic model has optimal policies which are obtainable from a deterministic model, then the qualitative behaviour of these policies can be derived by well-known arguments. Some of the main results are stated below without proof. The interested reader can refer to [7] for details or observe that if the budget constraints are used to eliminate the borrowing variables, then the deterministic model can be reduced to a liner production model so that these results follow from those of McFadden [9]. Next let $\bar{x}$ maximize $\{\rho E(F(x)) - x\}$ subject to $x \geqslant 0$. Assume $\bar{x}$ is unique and let $\bar{y} = E(F(\bar{x}))$. The special notation $\rho(\tau) = \sum_{t=1}^{\tau} \rho^{t-1}$, where ρ is a scalar, will be useful here and again in Section 5.

THEOREM 2. *The optimistic model has an optimal policy if and only if* $E(K) \geqslant 0$, *where*

$$K(\omega^{T+1}) = Y_0(\omega_1) + \sum_{t=1}^{T} \rho^{t-1}[\rho f_{\omega_{t+1}}(\bar{x}) - \bar{x}] - \rho^T Y_T(\omega^{T+1})$$

so that

$$E(K) = E(Y_0) + \rho(T)(\rho\bar{y} - \bar{x}) - \rho^T E(Y_T).$$

THEOREM 3. *Assume* $E(K) > 0$. *A feasible policy* (X_t, C_t, Z_t) *is optimal for the optimistic model if and only if there exist (nonrandom) competitive prices* $p_t \geqslant 0$ (*for* $t = 1,\dots, T+1$) *and* $q \geqslant 0$ *such that*

(i) c_t *maximizes* $\{\delta^{t-1}E(u(C)) - p_t E(C)\}$ *over all* $C(\omega^t) \geqslant 0$ *for* $t = 1,\dots, T$;

(ii) X_t *maximizes* $\{p_{t+1}E(F(X)) - p_t E(X)\}$ *over all* $X(\omega^t) \geqslant 0$ *for* $t = 1,\dots, T$;

(iii) Z_t *maximizes* $\{(\rho^{t-1}q - p_t)E(Z)\}$ *over all* $Z(\omega^t)$ *for* $t = 1,\dots, T+1$.

THEOREM 4. *Let* (X_t, C_t, Z_t) *be an optimal policy for the optimistic model. Then investment is stationary*, $X_t = \bar{x}$. *Moreover, consumption and borrowing are monotone*;

$$E(C_t) \geqslant E(C_{t+1}) \quad \textit{for} \quad t = 1,\dots, T-1,$$
$$E(Z_t) \leqslant E(Z_{t+1}) \quad \textit{for} \quad t = 2,\dots, T-1,$$

when $\rho > \delta$ (*the opposite inequalities holding when* $\rho < \delta$ *and equality holding when* $\rho = \delta$).

COROLLARY. *If* $\rho = \delta$ *then the optimal policy has the explicit form*

$$\begin{aligned} C_t &= \bar{c} && \textit{for} \quad t = 1,\dots, T, \\ Z_t &= Y_0(\omega^1) - \bar{x} - \bar{c} && \textit{for} \quad t = 1, \\ &= f_{\omega_t}(\bar{x}) - \bar{x} - \bar{c} && \textit{for} \quad t = 2,\dots, T, \\ &= \bar{y} - Y_T(\omega^T) && \textit{for} \quad t = T+1, \end{aligned}$$

where $\bar{c} = (\bar{y} - \bar{x}) + ((E(Y_0) - \bar{y}) + \rho^T(\bar{y} - E(Y_T)))/\rho(T)$.

4. THE PESSIMISTIC BORROWING MODEL

We now turn to the second version of the stochastic economic growth model with borrowing. Here the total borrowing must be repaid in every realization of the stochastic process. This change in the model results in a more complicated behaviour for the optimal policies.

The *pessimistic borrowing* model is defined by the following problem:

> Find random variables $X_t(\omega^t) \geqslant 0$, $C_t(\omega^t) \geqslant 0$ and $Z_t(\omega^t)$ which maximize $\sum_{t=1}^{T} \delta^{t-1} E(u(C_t))$ subject to the budget constraints
>
> $$X_1 + C_1 + Z_1 \leqslant Y_0,$$
>
> $$X_t + C_t + Z_t \leqslant F(X_{t-1}) \qquad \text{for} \quad t = 2,\ldots, T,$$
>
> $$Y_T + Z_{T+1} \leqslant F(X_T),$$
>
> and the repayment constraint
>
> $$\sum_{t=1}^{T+1} \rho^{t-1} Z_t \geqslant 0.$$

Again the constraints must hold pointwise (with respect to $\omega^t \in \Omega^t$). If for each ω^{T+1} the corresponding deterministic borrowing model is feasible, then there exists a feasible policy for the pessimistic model. The existence of optimal policies follows from familiar arguments in dynamic programming. We next turn to the duality theory for the pessimistic model.

DEFINITION. A feasible T-period policy (X_t, C_t, Z_t) from Y_0 to Y_T in the pessimistic model is *competitive* at prices $P_t(\omega^t) \geqslant 0$ (for $t = 1,\ldots, T+1$) and $Q(\omega^{T+1}) \geqslant 0$ if

(i) C_t maximizes $\{\delta^{t-1}E(u(C)) - E(P_t C)\}$ over all $C(\omega^t) \geqslant 0$, for $t = 1,\ldots, T$;

(ii) X_t maximizes $\{E(P_{t+1}F(X)) - E(P_t X)\}$ over all $X(\omega^t) \geqslant 0$, for $t = 1,\ldots, T$;

(iii) Z_t maximizes $\{E(E_{T-t+1}(\rho^{t-1}Q - P_t)Z)\}$ over all $Z(\omega^t)$, for $t = 1,\ldots, T+1$;

where the partial expectation E_{T-t+1} is defined (for $t = 1,\ldots, T+1$) by

$$E_{T-t+1}(W) = \int W(\omega_1,\ldots, \omega_{T+1})\, d\mathcal{P}^{T-t+1}(\omega_{t+1},\ldots, \omega_{T+1}).$$

These competitive conditions are similar to those for the optimistic model except that the prices are now random variables. Note that the partial expectation is an averaging over the randomness of a variable $W(\omega^{T+1})$ for only the remaining periods after time t. In particular, if $t = T+1$ then $E_0(W) = W$ and if we set $t = 0$, then $E_{T+1}(W) = E(W)$. The competitive conditions have a simple economic interpretation. Consumers maximize their utility subject to a budget constraint, producers maximize their profits and borrowers minimize their anticipated costs on their debts.

In order to circumvent a technical detail, we will assume that the repayment constraint is satisfied with equality in the optimal solutions. This can be achieved by increasing the final stock Y_T and avoids a separate treatment for the dual variable P_{T+1}.

THEOREM 5. *If (X_t^*, C_t^*, Z_t^*) is a competitive policy which satisfies the repayment constraint with equality in the pessimistic model then it is an optimal policy.*

Proof. Let (X_t, C_t, Z_t) be any feasible policy from Y_0 to Y_T, then by feasibility and the competitive conditions,

$$\begin{aligned}
&\sum_{t=1}^{T} \delta^{t-1}E(u(C_t)) - \sum_{t=1}^{T} \delta^{t-1}E(u(C_t^*)) \\
&\quad\leqslant \sum_{t=1}^{T} E(P_t(C_t - C_t^*)) \qquad \text{by (i)} \\
&\quad\leqslant \sum_{t=1}^{T} [(E(P_{t+1}F(X_t)) - E(P_tX_t)) - (E(P_{t+1}(F(X_t^*)) - E(P_tX_t^*))] \\
&\qquad - \sum_{t=1}^{T+1} E(P_t(Z_t - Z_t^*)) \\
&\quad\leqslant \sum_{t=1}^{T+1} E(E_{T-t+1}(\rho^{t-1}Q)(Z_t^* - Z_t)) \qquad \text{by (ii) and (iii)} \\
&\quad= \sum_{t=1}^{T+1} \int_{\Omega^t} \Big[(Z_t^*(\omega^t) - Z_t(\omega^t)) \\
&\qquad \times \int_{\Omega^{T-t+1}} \rho^{t-1}Q(\omega^{T+1})\, d\mathscr{P}^{T-t+1}(\omega_{t+1},\ldots,\omega_{T+1})\Big]\, dP^t(\omega^t) \\
&\quad= \int_{\Omega^{T+1}} \left[Q(\omega^{T+1}) \sum_{t=1}^{T+1} \rho^{t-1}(Z_t^*(\omega^t) - Z_t(\omega^t))\right] d\mathscr{P}^{T+1}(\omega^{T+1}) \\
&\quad\leqslant 0
\end{aligned}$$

since

$$Q \sum_{t=1}^{T+1} \rho^{t-1} Z_t^* = 0 \leqslant Q \sum_{t=1}^{T+1} \rho^{t-1} Z_t$$

holds for all $\omega^{T+1} \in \Omega^{T+1}$. ∎

Hence the existence of competitive prices is sufficient for optimality. To show necessity a Slater-type constraint qualification is imposed on the model. Specifically it is assumed that it is always possible to over-repay debts (perhaps by decreasing consumption).

THEOREM 6. *Let (X_t^*, C_t^*, Z_t^*) be an optimal T-period policy from Y_0 to Y_T for the pessimistic model. If there exists some feasible T-period policy from Y_0 to Y_T satisfying $\sum_{t=1}^{T+1} \rho^{t-1} Z_t(\omega^t) > 0$ for all $\omega^{T+1} \in \Omega^{T+1}$, then the policy (X_t^*, C_t^*, Z_t^*) is competitive.*

Proof. If the borrowing variables Z_t are eliminated by using the budget constraints, then it follows that (X_t^*, C_t^*) is optimal for the problem:

> Find $X_t \geqslant 0$, $C_t \geqslant 0$ which maximize $\sum_{t=1}^{T} \delta^{t-1} E(u(C_t))$
> subject to the (repayment) constraint
>
> $$\sum_{t=1}^{T} \rho^{t-1} C_t \leqslant Y_0 + \sum_{t=1}^{T} \rho^{t-1} (\rho F(X_t) - X_t) - \rho^T Y_T.$$

Letting $\mathscr{X}$ denote the policy (X_t, C_t), $\mathscr{F}(\mathscr{X})$ denote the objective $\sum_{t=1}^{T} \delta^{t-1} u(C_t)$ and $\mathscr{G}(\mathscr{X})$ denote the constraint function

$$Y_0 + \sum_{t=1}^{T} \rho^{t-1} (\rho F(X_t) - X_t) - \rho^T Y_T - \sum_{t=1}^{T} \rho^{t-1} C_t$$

then the pessimistic model is of the form:

> Find $\mathscr{X} \geqslant 0$ which maximizes $E\mathscr{F}(\mathscr{X})$ subject to $\mathscr{G}(\mathscr{X}) \geqslant 0$
> where $\mathscr{G}(\mathscr{X}) \in L_\infty(\Omega^{T+1})$.

The assumption that there exists some feasible policy which strictly satisfies the repayment constraint for all $\omega^{T+1} \in \Omega^{T+1}$ is equivalent to $\mathscr{G}(\mathscr{X}) > 0$ for some $\mathscr{X} \geqslant 0$. Hence by Luenberger [6, p. 224] there exists some $\psi^* \in L_\infty^*$ (the dual space of L_∞ with respect to the sup norm) such that $\psi^* \geqslant 0$ and $\mathscr{L}(\mathscr{X}, \psi^*)$ is maximized at $\mathscr{X}^* = (X_t^*, C_t^*)$ over all $\mathscr{X} \geqslant 0$ where the Lagrangian is given by

$$\mathscr{L}(\mathscr{X}, \psi) = E\mathscr{F}(\mathscr{X}) + \psi(\mathscr{G}(\mathscr{X})).$$

Now proceeding as in Zilcha [11] since $\psi^* \in \mathscr{L}_\infty^*$ there exists a bounded finitely additive measure η such that $\psi^*(\mathscr{G}) = \int \mathscr{G}\, d\eta$ for $\mathscr{G} \in L_\infty$. Moreover η has the decomposition $\eta = \eta_c + \eta_s$, where η_c is absolutely continuous with respect to $\mathscr{P}^{T+1}$ and η_s is finitely additive. Since η_c can be represented by a unique $Q \in \mathscr{L}_1(\Omega^{T+1})$, $Q \geqslant 0$, then we have

$$\begin{aligned}\psi^*(\mathscr{G}) &= \int Q \cdot \mathscr{G}\, d\mathscr{P}^{T+1} + \int \mathscr{G}\, d\eta_s \\ &= E(Q \cdot \mathscr{G}) + \int \mathscr{G}\, d\eta_s.\end{aligned}$$

Also there exists a sequence $\{A_n\}$ of subsets in $\mathscr{B}^{T+1}$, $A_n \subset A_{n+1}$ and $\eta_s(A_n) = 0$ for all n, with $\mathscr{P}^{T+1}(A_n) \to 1$ as $n \to \infty$. For $\mathscr{X} \in \mathscr{L}_\infty(\Omega^{T+1})$, $\mathscr{X} \geqslant 0$ define

$$\begin{aligned}\mathscr{X}^n(\omega^{T+1}) &= \mathscr{X}(\omega^{T+1}) && \text{if} \quad \omega^{T+1} \in A_n \\ &= \mathscr{X}^*(\omega^{T+1}) && \text{otherwise,}\end{aligned}$$

then

$$\begin{aligned}\mathscr{L}(\mathscr{X}^*, \psi^*) &= E\mathscr{F}(\mathscr{X}^*) + E(Q \cdot \mathscr{G}(\mathscr{X}^*)) + \int \mathscr{G}(\mathscr{X}^*)\, d\eta_s \\ &\geqslant E\mathscr{F}(\mathscr{X}^n) + E(Q \cdot \mathscr{G}(\mathscr{X}^n)) + \int \mathscr{G}(\mathscr{X}^*)\, d\eta_s\end{aligned}$$

holds for all n so that we have

$$\begin{aligned}&\int_{A_n} \mathscr{F}(\mathscr{X}^*)\, d\mathscr{P}^{T+1} + \int_{A_n} Q \cdot \mathscr{G}(\mathscr{X}^*)\, d\mathscr{P}^{T+1} \\ &\qquad \geqslant \int_{A_n} \mathscr{F}(\mathscr{X})\, d\mathscr{P}^{T+1} + \int_{A_n} Q \cdot \mathscr{G}(\mathscr{X})\, d\mathscr{P}^{T+1}.\end{aligned}$$

Next letting $n \to \infty$ we obtain

$$\begin{aligned}&\int \mathscr{F}(\mathscr{X}^*)\, d\mathscr{P}^{T+1} + \int Q \cdot \mathscr{G}(\mathscr{X}^*)\, d\mathscr{P}^{T+1} \\ &\qquad \geqslant \int \mathscr{F}(\mathscr{X})\, d\mathscr{P}^{T+1} + \int Q \cdot \mathscr{G}(\mathscr{X})\, d\mathscr{P}^{T+1}\end{aligned}$$

for all $\mathscr{X} \in \mathscr{L}_\infty(\Omega^{T+1})$, $\mathscr{X} \geqslant 0$. Reverting to the notation at the beginning of this proof, we have shown that

$$\sum_{t=1}^{T} \delta^{t-1} E(u(C_t)) + E\left(Q \cdot \left(Y_0 + \sum_{t=1}^{T} \rho^{t-1}(\rho F(X_t) - X_t) - \rho^T Y_T - \sum_{t=1}^{T} \rho^{t-1} C_t\right)\right)$$

is maximized at the policy (X_t^*, C_t^*) over all $X_t \geqslant 0$, $C_t \geqslant 0$. Since the random variables are non-anticipative, that is, $X_t(\omega^{T+1}) = X_t(\omega^t)$, $C_t(\omega^{T+1}) = C_t(\omega^t)$ and $\mathscr{P}^{T+1}$ is the product measure on the space $(\Omega^{T+1}, \mathscr{B}^{T+1})$, then upon defining random variables

$$P_t(\omega^t) = E_{T-t+1}(\rho^{t-1} Q(\omega^{T+1}))$$

the previous statement reduces to

$$\sum_{t=1}^{T} (\delta^{t-1} E(u(C_t)) - E(P_t C_t)) + \sum_{t=1}^{T} (E(P_{t+1} F(X_t) - E(P_t X_t)) + E(P_1 Y_0) - E(P_{T+1} Y_T))$$

being maximized at the policy (X_t^*, C_t^*) over all $X_t \geqslant 0$, $C_t \geqslant 0$. This defintion of prices P_t immediately yields competitive condition (iii). Conditions (i) and (ii) follow by separability of the variables in the above statement. ∎

As usual, the discussion of the qualitative behaviour of optimal policies is facilitated by the existence of competitive prices. In the case of the pessimistic borrowing model the utility function now plays a more significant role than in the deterministic (and hence optimistic) borrowing models.

THEOREM 7. *Assume that $u'(c)$ exists. If (X_t^*, C_t^*, Z_t^*) is an optimal policy for the pessimistic model and if $C_t^* > 0$, then*

$$E(u'(C_t^*)) \leqslant E(u'(C_{t+1}^*)) \qquad \text{for} \quad t = 1, \ldots, T-1$$

when $\rho > \delta$ (the opposite inequalities holding when $\rho > \delta$ and equality holding for $\rho = \delta$).

Proof. By differentiating condition (i) we have

$$\delta^{t-1} u'(C_t^*(\omega^t)) - P_t(\omega^t) \leqslant 0,$$

where equality holds if $C_t^*(\omega^t) > 0$. By condition (iii) $E_{T-t+1}(\rho^{t-1}Q(\omega^{T+1})) = P_t(\omega^t)$ holds for $t = 1,\ldots, T+1$. Thus if $E_1(P_{t+1})$ denotes expectation with respect to ω_{t+1} we have $E_1(P_{t+1}(\omega^{t+1})) = \rho P_t(\omega^t)$. Now assuming that $C_t^*(\omega^t) > 0$ holds for almost all ω^{T+1} and $t = 1,\ldots, T$, these equations combine to yield

$$\delta E_1(u'(C_{t+1}^*(\omega^{t+1}))) = \rho u'(C_t^*(\omega^t)).$$

Taking expected values (with respect to ω^t) yields the conclusions of the theorem. ▮

COROLLARY. *If u is a quadratic function, then* $E(C_t^*) \geqslant E(C_{t+1}^*)$ *for* $t = 1,\ldots, T-1$ *when* $\rho > \delta$ *(the opposite inequalities holding when* $\rho < \delta$ *and equality holding when* $\rho = \delta$*).*

Proof. Let $u'(C) = B - C$ where $B > 0$, then as $E(u'(C_t^*)) = B - E(C_t^*)$ the inequalities of the theorem statement are reversed. ▮

Theorem 7 together with Theorem 4 suggest strong similarities in the expected behaviour of the consumption policies of the two versions of the borrowing model. These similarities will be examined in more detail in Section 5. The investment policy for the optimistic model was stationary (at the "golden-rule" level $\bar{x}$). Investment in the pessimistic model, while not stationary, does belong to a set analogous to that described by Brock and Mirman [1]. Let $\bar{X}(\omega)$ maximize $\{\rho f_\omega(x) - x\}$ for all $x \geqslant 0$ (recall that f_ω is concave) and let $\bar{x}_m, \bar{x}_M$ denote the infimum and supremum values of $\bar{X}(\omega)$ on Ω.

THEOREM 8. *If* (X_t^*, C_t^*, Z_t^*) *is an optimal policy for the pessimistic model then* $\bar{x}_m \leqslant X_t^* \leqslant \bar{x}_M$ *for* $t = 1,\ldots, T$.

Proof. By condition (ii), $X_t^*(\omega^t)$ maximizes

$$\int P_{t+1}(\omega^{t+1}) f_{\omega_{t+1}}(X_t(\omega^t))\, d\mathscr{P}(\omega_{t+1}) - P_t(\omega^t)\, X_t(\omega^t).$$

Using conditions (i) and (iii) (as in the proof of the previous theorem) this is equivalent to, $X_t^*(\omega^t)$ maximizes

$$\frac{1}{\rho}\int P_{t+1}(\omega^{t+1})[\rho f_{\omega_{t+1}}(x) - x]\, d\mathscr{P}(\omega_{t+1}).$$

Since $P_{t+1}(\omega^{t+1}) \geqslant 0$ this then implies the conclusion of the theorem. ▮

Note that if $\bar{x}_M = \bar{x}_m$ then the optimal investment policy will be stationary and at the golden-rule level $\bar{x}$ (for $f = E(F)$).

5. COMPARISONS

The purpose of this final section is to compare the optimal consumption policies for the optimistic and pessimistic models. Specifically we show that the two policies are very similar during the initial periods of a model with a long planning horizon. Furthermore the pessimistic policy exhibits an adaptive behaviour with respect to the realized path of the production process.

Assume that an investment policy $X_t(\omega^t)$ has been chosen. Then the *total capital accumulation* is defined by

$$K_T(\omega^{T+1}) = Y_0(\omega_1) + \sum_{t=1}^{T} \rho^{t-1}[\rho f_{\omega_{t+1}}(X_t(\omega^t)) - X_t(\omega^t)] - Y_T(\omega^{T+1}).$$

Recall that an investment policy $X_t(\omega^t)$ is feasible for the optimistic model if $E(K_T) \geqslant 0$ or for the pessimistic model if $K_T(\omega^{T+1}) \geqslant 0$ for all ω^{T+1}.

If the investment policies are fixed, then the statements of the two models can be simplified (by elimination of the borrowing variables Z_t). A consumption policy $C_t \geqslant 0$ is *optimal* (with respect to the chosen X_t) for the optimistic model if it solves:

$$\begin{aligned} &\text{Maximize} && \sum_{t=1}^{T} \delta^{t-1} E(u(C_t)) \\ &\text{subject to} && \sum_{t=1}^{T} \rho^{t-1} E(C_t) \leqslant E(K_T). \end{aligned}$$

Similarly, a consumption policy $C_t \geqslant 0$ is *optimal* (with respect to the chosen X_t^*) for the pessimistic model if it solves:

$$\begin{aligned} &\text{Maximize} && \sum_{t=1}^{T} \delta^{t-1} E(u(C_t)) \\ &\text{subject to} && \sum_{t=1}^{T} \rho^{t-1} C_t(\omega^t) \leqslant K_T^*(\omega^{T+1}) \\ &\text{for all} && \omega^{T+1} \in \Omega^{T+1}. \end{aligned}$$

But by the previous sections, optimal policies exist and can be assumed to be competitive. Thus there exist prices q and Q for the two models respectively such that the following constraints are satisfied:

$$\sum_{t=1}^{T} \rho^{t-1} E(C_t) = \sum_{t=1}^{T} \rho^{t-1} v((\rho/\delta)^{t-1} q) = E(K_T),$$

$$\sum_{t=1}^{T} \rho^{t-1} C_t = \sum_{t=1}^{T} \rho^{t-1} v((\rho/\delta)^{t-1} E_{T-t+1}(Q)) = K_T^*(\omega^T),$$

where v is the inverse function of the marginal utility function u' and $K_T^*(\omega^T)$ is the minimum of $K_T^*(\omega^{T+1})$ as a function of ω_{T+1}.

In order to carry out our comparison of optimal consumption policies we need to solve these constraints for the competitive prices q and Q. Explicit solutions can be obtained for the case of a quadratic utility function u. In this case $v(q) = B - q$, where $B > 0$ is the "bliss point" of u.

Since the calculations for the pessimistic model are somewhat complicated, the following notation will be helpful. Let $\nu = \rho^2/\delta$, and define

$$\rho(\tau) = \sum_{t=1}^{\tau} \rho^{t-1},$$

$$\nu(\tau) = \sum_{t=1}^{\tau} \nu^{t-1},$$

$$W_T = (\rho(\tau)B - K_T)/\nu(T),$$

$$W_T^* = (\rho(\tau)B - K_T^*)/\nu(T).$$

The subscript T will usually be omitted in W_T and W_T^*. Some simple identities for $\nu(\tau)$ will be needed (their proofs, being straightforward manipulations, are omitted).

LEMMA 1. *For any $\nu > 0$ and integer $T > 0$,*

(i) $\nu(T) - \nu(t-1) = \nu^{t-1}\nu(T-t+1)$ *for* $t = 1, \ldots, T$,

(ii) $$\sum_{t=1}^{T} \frac{\nu^{t-1}}{\nu(t)\,\nu(t+1)} = 1 - \frac{\nu^T}{\nu(T+1)},$$

(iii) $$\sum_{s=1}^{t-1} \frac{\nu^{s-1}\nu(T)}{\nu(T-t+s)\,\nu(T-t+s+1)} = \frac{\nu(t-1)}{\nu(T-t+1)} \quad \textit{for} \quad t = 2, \ldots, T.$$

The prices for the optimistic and pessimistic models are defined by the expressions

$$q = E(W)$$

and

$$Q(\omega^{T+1}) = W^* - \sum_{s=1}^{T-1} \nu^{s-T}\left[\frac{\nu(T)}{\nu(s)\,\nu(s+1)} E_s(W^*) - W^*\right].$$

Note that as $W^* = W^*(\omega^T)$ then Q is actually independent of ω_{T+1}. Thus we have the equation $E_{T-t+1}(Q(\omega^{T+1})) = E_{T-t}(Q(\omega^T))$ for $t = 1, 2, \ldots, T$. Note that $E_0(A) = A$ for any random variable A.

LEMMA 2. *$E_{T-1}(Q) = E_{T-1}(W^*)$ and for $t = 2, \ldots, T$*

$$E_{T-t}(Q) = \frac{\nu(T)}{\nu^{t-1}}\left[\frac{E_{T-t}(W^*)}{\nu(T-t+1)} - \sum_{s=1}^{t-1} \frac{\nu^{s-1}E_{T-t+s}(W^*)}{\nu(T-t+s)\,\nu(T-t+s+1)}\right].$$

Proof. *Applying E_{T-t} to the expression for $Q(\omega^T)$ we have, upon collecting common terms*

$$E_{T-t}(Q) = E_{T-t}(W^*) - \sum_{s=1}^{T-1} \nu^{s-T}\left[\frac{\nu(T)}{\nu(s)\,\nu(s+1)} E_{T-t}(E_s(W^*)) - E_{T-t}(W^*)\right]$$

$$= \frac{\nu(T)}{\nu^{T-1}}\left[E_{T-t}(W^*) - \sum_{s=1}^{T-1} \frac{\nu^{s-1}}{\nu(s)\,\nu(s+1)} E_{T-t}(E_s(W^*))\right].$$

But $E_{T-t}(E_s(W^*))$ equals $E_{T-t}(W^*)$ for $s = 1, \ldots, T-t$ and equals $E_s(W^*)$ for $s = T-t+1, \ldots, T-1$. This second possibility does not arise for the case $t = 1$. When $t = 1$, we have

$$E_{T-1}(Q) = \frac{\nu(T)}{\nu^{T-1}}\left[1 - \sum_{s=1}^{T-1} \frac{\nu^{s-1}}{\nu(s)\,\nu(s+1)}\right] E_{T-1}(W^*) = E_{T-1}(W^*)$$

by identity (ii) of Lemma 1. For $t = 2, \ldots, T$,

$$E_{T-t}(Q) = \frac{\nu(T)}{\nu^{T-1}}\left[\frac{\nu^{T-t}}{\nu(T-t+1)} E_{T-t}(W^*) - \nu^{T-t}\sum_{s=1}^{t-1} \frac{\nu^{s-1}E_{T-t+s}(W^*)}{\nu(T-t+s)\,\nu(T-t+s+1)}\right],$$

which simplifies to the required form upon applying the same identity. ∎

Lemma 3. *The prices q and Q are competitive for the optimistic and pessimistic models.*

Proof. In the optimistic model:

$$\begin{aligned}\sum_{t=1}^{T} \rho^{t-1}v((\rho/\delta)^{t-1}q) &= \rho(T)B - \nu(T)q\\ &= \rho(T)B - \nu(T)\,E(W)\\ &= \rho(T)B - \nu(T)[(\rho(T)B - E(K_T))/\nu(T)]\\ &= E(K_T)\end{aligned}$$

as required. In the pessimistic model, applying Lemma 2,

$$\begin{aligned}&\sum_{t=1}^{T} \nu^{t-1}E_{T-t}(Q(\omega^T))\\ &\quad= E_{T-1}(W^*) + \nu(T)\sum_{t=2}^{T}\frac{E_{T-t}(W^*)}{\nu(T-t+1)}\\ &\quad- \nu(T)\sum_{t=2}^{T}\sum_{s=1}^{t-1}\frac{\nu^{s-1}E_{T-t+s}(W^*)}{\nu(T-t+s)\,\nu(T-t+s+1)}.\end{aligned}$$

The first term is combined with the second, the double summation is rearranged (using the change of variable $r = t - s$), to obtain

$$\begin{aligned}&= \nu(T)\left[\sum_{t=1}^{T}\frac{E_{T-t}(W^*)}{\nu(T-t+1)} - \sum_{r=1}^{T-1}\sum_{s=1}^{T-r}\frac{\nu^{s-1}E_{T-r}(W^*)}{\nu(T-r)\,\nu(T-r+1)}\right]\\ &= \nu(T)\left[\sum_{t=1}^{T}\frac{E_{T-t}(W^*)}{\nu(T-t+1)} - \sum_{r=1}^{T-1}\frac{E_{T-r}(W^*)}{\nu(T-r+1)}\right]\\ &= \nu(T)\,E_0(W^*)/\nu(1) = \nu(T)\,W^*.\end{aligned}$$

Hence

$$\begin{aligned}&\sum_{t=1}^{T} \rho^{t-1}v((\rho/\delta)^{t-1}\,E_{T-t+1}(Q(\omega^{T+1}))\\ &\quad= \rho(T)B - \sum_{t=1}^{T}\nu^{t-1}E_{T-t}(Q(\omega^T))\\ &\quad= \rho(T)B - \nu(T)\,W^*\\ &\quad= \rho(T)B - (\rho(T)B - K_T^*) = K_T^*\end{aligned}$$

as required. ∎

THEOREM 9. *If u is a quadratic utility function (with bliss point B), then the optimal consumption policies are*

$$C_t^T = B - (\rho/\delta)^{t-1} E(W_T)$$

in the optimistic model and

$$C_t^T = B - (\rho/\delta)^{t-1} E_{T-t}(W_T^*) + D_t^T$$

in the pessimistic model, where the deviation $D_1^T = 0$ *and*

$$D_t^T = \rho^{1-t} \sum_{s=1}^{t-1} \nu^{s-1} \left[\frac{\nu(T) E_{T-t+s}(W_T^*)}{\nu(T-t+s)\,\nu(T-t+s+1)} - \frac{E_{T-t}(W_T^*)}{\nu(T-t+1)} \right]$$

for $t = 2,\dots, T$.

Proof. In the optimistic model

$$C_t^T = v((\rho/\delta)^{t-1} q) = B - (\rho/\delta)^{t-1} E(W)$$

as claimed. In the pessimistic model

$$C_t^T = v((\rho/\delta)^{t-1} E_{T-t}(Q(\omega^T)))$$

and applying Lemma 2

$$C_1^T = B - E_{T-1}(Q) = B - E_{T-1}(W^*);$$

and for $t > 1$,

$$\begin{aligned}
C_t^T &= B - (\rho/\delta)^{t-1} E_{T-t}(Q) \\
&= B - (\rho/\delta)^{t-1} \frac{\nu(T)}{\nu^{T-1}} \left[\frac{E_{T-t}(W^*)}{\nu(T-t+1)} \right. \\
&\qquad \left. - \sum_{s=1}^{t-1} \frac{\nu^{s-1} E_{T-t+s}(W^*)}{\nu(T-t+s)\,\nu(T-t+s+1)} \right] \\
&= B - (\rho/\delta)^{t-1} E_{T-t}(W^*) \\
&\qquad + \frac{1}{\rho^{t-1}} \left[\left(\nu^{t-1} - \frac{\nu(T)}{\nu(T-t+1)} \right) E_{T-t}(W^*) \right. \\
&\qquad \left. + \nu(T) \sum_{s=1}^{t-1} \frac{\nu^{s-1} E_{T-t+s}(W^*)}{\nu(T-t+s)\,\nu(T-t+s+1)} \right],
\end{aligned}$$

which reduces to the required expression upon applying identity (i) of Lemma 1. ∎

COROLLARY. *$E(D_t^T) = 0$ for $t = 1, \dots, T$.*

Proof. This clearly holds for $t = 1$. If $t > 1$ then

$$\begin{aligned} E(D_t^T) &= \rho^{1-t} \sum_{s=1}^{t-1} \nu^{s-1} \left[\frac{\nu(T)}{\nu(T-t+s)\,\nu(T-t+s+1)} \right. \\ &\quad \left. - \frac{1}{\nu(T-t+1)} \right] E(W^*) \\ &= \rho^{1-t} \left[\left(\sum_{s=1}^{t-1} \frac{\nu^{s-1}}{\nu(T-t+s)\,\nu(T-t+s+1)} \right) \nu(T) \right. \\ &\quad \left. - \frac{\nu(T-1)}{\nu(T-t+1)} \right] E(W^*) \\ &= 0 \end{aligned}$$

by identity (iii) of Lemma 1. ∎

Although the consumption policies of the theorem have similar forms, the optimistic policy is based upon the expected capital accumulation over the entire horizon whereas the pessimistic policy depends upon the realized past history up until time t and the expected capital accumulation over the remaining horizon. The corollary shows that the deviations in the pessimistic consumption policy do average out to be zero (the deviation in the initial period being exactly zero).

For certain choices of ρ and δ the repayment constraint becomes less important as the planning horizon T tends towards infinity. Our final results show that for such choices the distinction between the optimistic and pessimistic models disappears.

THEOREM 10. *Let u be a quadratic utility (with bliss point B). If $\rho < 1$ and $\nu \geqslant 1$ then $\lim_{T \to \infty} D_t^T = 0$ for each t.*

Proof. Since $\rho < 1$ and $f_\omega(x) - x < M$ for all x, all ω, then it follows that $K_T^*(\omega^T)$ is uniformly bounded (for reasonable initial and final stocks Y_0, Y_T). But as

$$W_T^* = (\rho(T)B - K_T^*)/\nu(T)$$

it follows that $W_T^*(\omega^T)$ is uniformly bounded in ω^T and in T. Now if $\nu \geqslant 1$ it follows that

$$\frac{1}{\nu(T-t+1)} E_{T-t}(W_T^*)$$

and

$$\frac{\nu(T)}{\nu(T-t+s)\,\nu(T-t+s+1)}\,E_{T-t+s}(W_T^*)$$

tend to zero (for fixed t, s) as $T \to \infty$ so that $\lim_{T\to\infty} D_t^T = \rho^{1-t} \sum_{s=1}^{t-1} \nu^{s-1} \cdot 0 = 0$. ∎

Corollary. *If $\rho < 1$ and $\nu \geqslant 1$ then the optimal consumption policies in both models satisfy $\lim_{T\to\infty} C_t^T = B$ or equivalently $\lim_{T\to\infty} P_t^T = 0$ for each finite t.*

Proof. From Theorems 9 and 10 we have (for the pessimistic model) $\lim_{T\to\infty} C_t^T = \lim_{T\to\infty}(B - (\rho/\delta)^{t-1} E_{T-t}(W_T^*) + D_t^T) = B$ since $W_T^* \to 0$ as $T \to \infty$ under the conditions on ρ and ν. $P_t^T \to 0$ follows from $P_t^T = \delta^{t-1}u'(C_t^T) = \delta^{t-1}(B - C_t^T)$. ∎

ACKNOWLEDGMENTS

The authors would like to express their thanks to several referees for helpful suggestions on an earlier version of this paper.

REFERENCES

1. W. A. Brock and L. J. Mirman, Optimal economic growth and uncertainty: The discounted case, *J. Econ. Theory* **4** (1972), 479–513.
2. D. Cass, Optimal growth in an aggregative model of capital accumulation, *Rev. Econ. Studies* **32** (1965), 233–240.
3. D. Gale, A mathematical theory of optimal economic development, *Bull. Amer. Math. Soc.* **74** (1968), 207–223.
4. N. H. Hakansson, Optimal investment and consumption strategies under risk for a class of utility functions, *Econometrica* **38** (1970), 587–607.
5. T. C. Koopmans, On the concept of optimal economic growth, *Pontificau Acad. Sci. Scripta Varia* **28** (1965), 225–300.
6. D. G. Luenberger, "Optimization by Vector Space Methods," Wiley, New York, 1969.
7. L. C. MacLean, "Optimal Economic Growth," doctoral dissertation, Department of Mathematics, Dalhousie University, Halifax, N.S., May 1976.
8. L. C. MacLean, C. A. Field, and W. R. S. Sutherland, "Optimal Growth with Borrowing," Research Report No. 22, Mathematics Department, Dalhousie University, August 1979.
9. D. McFadden, The evaluation of development programmes, *Rev. Econ. Studies* **34** (1967), 25–50.
10. M. E. Yaari, Uncertain lifetime, life insurance and the theory of the consumer, *Rev. Econ. Studies* **32** (1965), 137–150.
11. I. Zilcha, On competitive prices in a multisector economy with stochastic production and resources, *Rev. Econ. Studies* **43** (1976), 431–438.

JOURNAL OF ECONOMIC THEORY **24**, 187–204 (1981)

Noncooperative Oligopoly with Entry*

KOFI O. NTI

School of Business, Washington University, St. Louis, Missouri 63130

AND

MARTIN SHUBIK

Cowles Foundation for Research in Economics,
Yale University, New Haven, Connecticut 06520

Received May 21, 1979; revised April 7, 1980

1. INTRODUCTION

The literature on oligopolistic competition with entry is substantial (Nti and Shubik [11]). However, with the exception of Shubik [13], Friedman [4], Spence [14], Dubey and Shubik [2], and a few others, most of the theoretical analyses on entry tend to ignore strategic issues. The well-known results of Kamien and Schwartz [7], Gaskins [6], and others, have generally focussed on how a group of oligopolists may collectively control entry by their pricing policies. But established firms need not act collectively. They may engage in intra-oligopolistic competition and, above all, they may perceive entry threats differently. In addition, potential entrants in oligopolistic industries are often big firms whose interests and strategic options need to be explicitly considered. Hence, strategic issues, involving varying degrees of cooperation and antagonism, are critical to the entry problem.

Furthermore, strategic oligopoly with entry is pertinent to research on Cournot–Walras equilibrium; it provides an entry-inclusive partial equilibrium building block to supplement general equilibrium models that integrate oligopolistic production and competitive exchange markets (Gabszewicz and Vial [5], Novshek and Sonnenschein [9]).

In this paper, competition between several established firms in an oligopolistic industry with a finite number of potential entrants is modelled as a noncooperative game with quantity and entry decisions as strategic variables. The varieties of disadvantages facing potential entrants are

---

* Partially supported by the Office of Naval Research under contract N0014-77-C-0518. Parts of this paper appeared in Dr. Nti's Ph. D. thesis, Yale University, 1977.

0022-0531/81/020187–18$02.00/0

Copyright © 1981 by Academic Press, Inc.

aggregated into single period fixed cost barriers. The existence of a noncooperative equilibrium solution is proved. At equilibrium, no active firm has an incentive to change production or exit and no firm outside the industry has an incentive to enter. Apart from potential entrants who may randomize their entry decisions, the equilibrium solution involves no randomization in the outputs of any firm. The solutions are characterized and their dependence on entry costs are determined. The solutions partition entry costs into several regions which have natural interpretations regarding the ease or difficulty of entry and the use of threats or exclusionary tactics to enforce certain outcomes.

The model is formulated in Section 2. Cournot oligopoly without entry is analyzed in Section 3. And in Sections 4 and 5, the main results proving and characterizing the noncooperative equilibrium solution for the game with one potential entrant are established and discussed. Section 6 extends the results to the situation with several asymmetric potential entrants. We conclude with a brief review and discussion of the results of the paper.

2. THE MODEL

Consider a one product oligopolistic industry with n established firms and k potential entrants. The sole decision variable of an established firm is the quantity of the product it should offer for sale. A potential entrant, on the other hand, must simultaneously decide whether it should actually enter into competition and the quantity of the product it should offer for sale. All decisions are made independently.

We study a highly symmetric model as this serves to illustrate the qualitative results we wish to show with analytical simplicity.

The price, p, of the commodity is determined by the relation

$$p = \phi(q), \tag{1}$$

where q is the total quantity of the commodity offered for sale. All firms, established as well as potential entrants, face identical variable cost function $c(q_i)$, where q_i is the output of the ith firm. Established firms have no fixed costs but each potential entrant incurs a non-negative fixed cost D if it actually enters into competition and offers any quantity of the product for sale. The objective is to identify and characterize noncooperative equilibrium solutions for the problem under the following assumptions:

ASSUMPTION 1. *The variable cost function $c(q_i)$ is twice differentiable with*

$$c'(q_i) > 0, c''(q_i) \geqslant 0. \tag{2}$$

ASSUMPTION 2. *For all total outputs, q, offered for sale by all market participants, the demand function $\phi(q)$ is decreasing for $0 \leqslant q \leqslant Q$. Also $\phi(0) > 0$ and there exists Q for which $\phi(Q) = 0$ and $\phi(q) = 0$ for all $q \geqslant Q$. Furthermore, $\phi(q)$ is twice differentiable and satisfies*

$$\phi'(q) < 0, \phi''(q) \leqslant 0 \qquad \text{for all} \quad 0 \leqslant q \leqslant Q. \tag{3}$$

We also assume that all firms have equal capacity and that the firms in the industry have enough capacity to supply the Cournot equilibrium requirements. This avoids having to discuss boundary solutions.

Assumption 1 states that each firm's variable cost is a non-decreasing function of its output and its marginal cost exhibits increasing returns. Assumption 2 describes a negatively sloping demand function with decreasing first derivative. These two assumptions are typical of cost and demand functions in the theory of the firm. Together, they imply concave profit functions and help establish the existence of noncooperative equilibrium for classical oligopoly without potential entrants. In the framework of symmetric oligopoly, following Roberts and Sonnenschein [12], it is possible to dispense with implied assumptions of concavity (and differentiability) but to do so here would only raise technical complexities in subsequent optimizations, especially when we attempt to characterize equilibrium solutions for the problem with entry.

Before considering the entry problem, we need some results for the Cournot oligopoly problem without potential entrants.

3. COURNOT OLIGOPOLY WITHOUT POTENTIAL ENTRANTS

Consider n identical oligopolists competing with quantity as strategic variables under the assumptions of the model.

Suppose the ith firm offers an output q_i for sale, $i = 1,..., n$. Then the profit, π_j, of the jth firm is given by

$$\pi_j = q_j \phi\left(\sum_{i=1}^{n} q_i\right) - c(q_j), \qquad j = 1, 2,..., n. \tag{4}$$

In general, a strategy for the ith firm may involve randomization over all non-negative outputs. A noncooperative equilibrium solution will then be a set of n strategies, one for each firm, with the property that no firm acting alone can improve its expected profit if it uses a strategy other than its part of the equilibrium solution.

The following theorem asserts that without loss of generality we may restrict our search for noncooperative equilibrium solutions to pure strategies. That is, no firm needs to randomize its production decision and we can look for an n-tuple of output $(q_1^*, q_2^*, \ldots, q_n^*)$ such that

$$\pi_j^* = q_j^* \phi\left(\sum_{i=1}^{n} q_i^*\right) - c(q_j^*)$$
$$\geqslant q_j \phi\left(\sum_{i \neq j}^{n} q_i^* + q_j\right) - c(q_j), \quad \text{for all } q_j \geqslant 0,\ j = 1, 2, \ldots, n. \tag{5}$$

To establish the existence of a n-tuple of noncooperative equilibrium production levels we need a theorem of Nikaido and Isoda [8].

THEOREM 1 (Nikaido and Isoda). *If the following conditions hold for an n-player game*

(i) *The set of decisions, Σ_i, for each player is a bounded, closed and convex subset of Euclidean space R^k.*

(ii) *The payoff function $\pi_i(q_1, \ldots, q_n)$ is concave in q_i, for fixed $q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n$, $i = 1, 2, \ldots, n$, $q_i \in \Sigma_i$.*

(iii) *$\pi_i(q_1, \ldots, q_n)$ is continuous in $q_1, \ldots, q_n$.*

Then there exists an n-tuple $(q_1^*, \ldots, q_n^*)$ with $q_i^* \in \sum_i$, $i = 1, \ldots, n$ which is a noncooperative equilibrium solution for the game.

Under the assumptions of the Cournot oligopoly problem above, it is clear no firm needs to produce more than Q units of the commodity. Hence condition (i) is satisfied. Also the second partial derivative of π_i with respect to q_i is nonpositive for all i so the concavity requirements in condition (ii) are met. It is also clear π_i is continuous in all variables. Hence:

COROLLARY 1. *Under the assumptions of the model, there exists a noncooperative equilibrium solution of the form $(q_1^*, \ldots, q_n^*)$ with $0 \leqslant q_i^* \leqslant Q$, $i = 1, 2, \ldots, n$, to the Cournot oligopoly problem.*

A characterization of the pure strategy noncooperative equilibrium solution, which we state as a theorem, is given in Burger [1, pp. 49–52].

THEOREM 2. *Under assumptions 1 and 2 the n-firm Cournot oligopoly problem has exactly one noncooperative equilibrium point $(q^*, \ldots, q^*)$ where*

(a) *$q^* = 0$ if $c'(0) \geqslant \phi(0)$ or*

(b) *q^* is the only root of the equation*

$$H_n(q) = \phi(nq) + q\phi'(nq) - c'(q) = 0 \tag{6}$$

on the interval $0 < q < Q/n$ if $c'(0) < \phi(0)$.

The condition $c'(0) < \phi(0)$, for which there exists a positive output for each firm at the noncooperative equilibrium solution, only says that marginal cost at zero output is less than the market price. This is obviously a necessary condition for a firm to send a positive output to the market. On the other hand. if marginal cost at zero output exceeds the market price, then it is unprofitable for any firm to produce. From henceforth, we will rule out this trivial situation and assume each active firm has an incentive to send a positive output to the market.

ASSUMPTION 3. *The functions $\phi(\cdot)$ and $c(\cdot)$ satisfy $\phi(0) > c'(0)$.*

The noncooperative equilibrium solutions to the Cournot oligopoly problem may be parametrized with respect to the number of active firms.

Let

$$H_n(\xi) = \phi(n\xi) + \xi\phi'(n\xi) - c'(\xi). \tag{7}$$

Then $H_n'(\xi) < 0$ for $0 < \xi < Q/n$. And the unique noncooperative equilibrium output, ξ_n, for each firm in the n-firm oligopoly game satisfies

$$H_n(\xi_n) = 0 \qquad \text{with} \qquad 0 < \xi_n < Q/n. \tag{8}$$

It is also straightforward to establish

THEOREM 3. *For $m > n$, let ξ_m and ξ_n be the noncooperative equilibrium outputs for the m-firm and n-firm Cournot oligopolies respectively. Also let π^m and π^n be the respective profits of the firms.*

Then

(a) $\xi_n > \xi_m$,

(b) $\pi^n > \pi^m$,

(c) $n\xi_n < m\xi_m$.

Theorem 3 states that the individual firms in the n-firm oligopoly problem produce more and earn more than their counterparts in the m-firm problem but the total output in the n-firm problem is less than the total output in the m-firm problem.

4. COURNOT OLIGOPOLY WITH ONE POTENTIAL ENTRANT

Continuing with the entry model we now assume there are n established firms and one potential entrant. The potential entrant is designated the 0th firm and the established firms are labelled from 1 through n.

Define δ by

$$\begin{aligned} \delta &= 0 \qquad \textit{if potential entrant stays out.} \\ &= 1 \qquad \textit{if potential entrant enters.} \end{aligned}$$

Then the profits may in general be written as

$$\pi_j^\delta = q_j \phi \left(\sum_{l=1}^{n} q_l + \delta q_0 \right) - c(q_j), \qquad j = 1,\ldots, n, \tag{9}$$

and

$$\pi_0^\delta = \delta \left[q_0 \phi \left(\sum_{l=1}^{n} q_l + \delta q_0 \right) - c(q_0) - D \right]. \tag{10}$$

A strategy for the jth active firm is a distribution function $F_j(\xi)$ over all nonnegative outputs, where $F_j(\xi) = \mathrm{Prob}\{q_j \leqslant \xi\}$, $j = 1,\ldots, n$. Similarly, a strategy for the potential entrant is a probability distribution $F_0(\xi)$ over all nonnegative outputs if it enters and a discrete probability distribution h over its entry decision, where

$$F_0(\xi) = \mathrm{Prob}\{q_0 \leqslant \xi \mid \delta = 1\}$$

and

$$h_\delta = \mathrm{Prob}\{\delta = \nu\}, \qquad \nu = 0, 1.$$

The expected profit of the firms under the strategies F_0, F_j, h_δ are

$$E\pi_j^\delta = \sum_{\delta=0}^{1} \left[\int \prod_{l=1}^{n} dF_l(q_l) \pi_0^\delta \right] h_\delta, \qquad j = 1,\ldots, n, \tag{11}$$

and

$$E\pi_0^\delta = \sum_{\delta=0}^{1} \left[\int \prod_{l=1}^{n} dF_l(q_l) \pi_0^\delta \right] h_\delta. \tag{12}$$

A noncooperative equilibrium strategy for the problem is, in general, a set of $n + 2$ probability distributions $\{F_j^*\}$, $j = 0,\ldots, n$, h_δ^* satisfying

$$\begin{aligned} &\sum_{\delta=0}^{1} \int \left[\prod_{l=1}^{n} dF_l^*(q_l) \pi_j^\delta \right] h_\delta^* \\ &\qquad \geqslant \sum_{\delta=0}^{1} \int \left[\prod_{\substack{l=1 \\ l \neq j}}^{n} dF_l^*(q_l)\, dF_j(q_j) \pi_j^\delta \right] h_\delta^* \end{aligned} \tag{13}$$

for $j = 1,\dots, n$, and all distribution functions F_j. And

$$\sum_{\delta=0}^{1} \left[\int \prod_{l=1}^{n} dF_l^*(q_l)\, dF_0^*(q_0)\pi_0^\delta\right] h_\delta^*$$
$$\geqslant \sum_{\delta=0}^{1} \left[\int \prod_{l=1}^{n} dF_l^*(q_l)\, dF_0(q_0)\pi_0^\delta\right] h_\delta \tag{14}$$

for all distribution functions F_0 and probability distribution h_δ.

We will show that without loss of generality, the search for noncooperative equilibrium solutions may be restricted to pure production strategies for all firms and a possible randomization in the entry decision of the potential entrant.

That is, if we let δ be the probability of entry with $0 \leqslant \delta \leqslant 1$ and compute the expected profit from (9) and (10) as follows:

$$\pi_0^\delta = \delta \left[q_0 \phi\left(\sum_{l=1}^{n} q_l + q_0\right) - c(q_0) - D\right] \tag{15}$$

and

$$\pi_j^\delta = (1-\delta)\left[q_j \phi\left(\sum_{l=1}^{n} q_l\right) - c(q_j)\right]$$
$$+ \delta\left[q_j \phi\left(\sum_{l=1}^{n} q_l + q_0\right) - c(q_j)\right], \qquad j = 1,\dots, n. \tag{16}$$

Then there exists an $n+2$-tuple $(q_1^*,\dots, q_n^*, q_0^*, \delta^*)$ with $0 \leqslant \delta^* \leqslant 1$ and $0 \leqslant q_i^* \leqslant Q$, $i = 0, 1,\dots, n$, such that

$$\pi_j^{\delta *}(q_1^*,\dots, q_{j-1}^*, q_j^*, q_{j+1}^*,\dots, q_n^*, q_0^*)$$
$$\geqslant \pi_j^{\delta *}(q_1^*,\dots, q_{j-1}^*, q_j, q_{j+1}^* \cdots q_n^*, q_0^*) \qquad \text{for all} \quad q_j,\ j = 1,\dots, n \tag{17}$$

and

$$\pi_0^{\delta *}(q_1^*,\dots, q_n^*, q_0^*) \geqslant \pi_0^\delta(q_1^*,\dots, q_n^*, q_0) \qquad \text{for all} \quad \delta \text{ and } q_0. \tag{18}$$

Observations. (a) The function π_j^δ and π_0^δ are continuous in δ, $q_0, q_1,\dots, q_n$.

(b) The feasible decision set for each firm is convex and compact, $0 \leqslant \delta \leqslant 1$, $0 \leqslant q_l \leqslant Q$, $j = 0,\dots, n$.

(c) For fixed δ, $q_0, q_1,\dots, q_{j-1}, q_{j+1},\dots, q_n$ the function π_j^δ is concave in q_j, $j = 1,\dots, n$.

(d) For fixed $q_1,\dots, q_n$, the function π_0^δ is concave *separately* in q_0 and δ but not *jointly* in δ and q_0.

For fixed $\delta, q_0, q_1, \ldots, q_{j-1}, q_{j+1}, \ldots, q_n$ let

$$B_j(\delta, q_{-j}) = \{q_j^* \mid \pi_j^\delta(q_0, \ldots, q_j^*, \ldots, q_n) \geqslant \pi_j^\delta(q_0, \ldots, q_j, \ldots, q_n)\} \quad \text{for } j = 1, 2, \ldots, n. \tag{19}$$

And for fixed $q_1, q_2, \ldots, q_n$ *let*

$$B_0(q_{-0}) = \{(\delta^*, q_0^*) \mid \pi_0^{\delta^*}(q_0^*, q_1, \ldots, q_n) \geqslant \pi_0^\delta(q_0, \ldots, q_1, \ldots, q_n)\} \tag{20}$$

Observations. (e) For fixed $\delta, q_0, q_1, \ldots, q_{j-1}, q_{j+1}, \ldots, q_n$, $B_j(\delta, q_{-j})$ is the set of optimal responses for the jth firm. The set is nonempty since $\pi_j^\delta(q_0, q_1, \ldots, q_n)$ is concave in q_j and the decision set is closed and convex. Actually the second order differentiability conditions on the cost and demand functions $c(\cdot)$ and $\phi(\cdot)$ imply $\pi_j^\delta(q_0, q_1, \ldots, q_n)$ is strictly concave in q_j and so the jth firm has a unique optimal response.

(f) The set $B_0(q_{-0})$ is the set of optimal responses for the potential entrant for fixed $q_1, \ldots, q_n$. The set $B_0(q_{-0})$ is nonempty, not necessarily convex but it is contractible.

Let $I = \{t \mid 0 \leqslant t \leqslant 1\}$.

DEFINITION. A subset Z of Euclidean space R^n is *contractible* (deformable) into a point $z_0 \in Z$, if there exists a continuous function $h(t, z)$ mapping $I \times Z$ into Z such that for all $z \in Z$, $h(0, z) = z$ and $h(1, z) = z_0$.

We list without proof some standard results on contractible sets.[1]

(i) If a set Z is contractible to $z_0 \in Z$, then Z is contractible on any other point $z \in Z$.

(ii) The Cartesian product of a finite number of contractible sets in contractible.

(iii) A convex set is contractible.

THEOREM 4. *The set* $B_0(q_{-0})$ *is contractible.*

Proof. For fixed $q_1, q_2, \ldots, q_n$, the set $B_0(q_{-0})$ is defined by

$$\begin{aligned} B_0(q_{-0}) = \Bigg\{(\delta^*, q_0^*) \mid \delta^* \left[q_0^* \phi\left(\sum_{i=1}^{n} q_i + q_0^*\right) - c(q_0) - D\right] \\ = \max_{\substack{0 \leqslant \delta \leqslant 1 \\ 0 \leqslant q_0 \leqslant Q}} \delta \left[q_0 \phi\left(\sum_{i=1}^{n} q_i + q_0\right) - c(q_0) - D\right]\Bigg\}. \end{aligned} \tag{21}$$

---

[1] See, for example, Murray Eisenberg, "Topology" p. 372, Holt, Rinehart & Winston, New York: 1974.

Since $q_0\phi(\sum_{i=1}^n q_i + q_0) - c(q_0) - D$ is (strictly) convex in q_0, there exists a (unique) point $\bar{q}_0$ satisfying

$$\bar{q}_0\phi\left(\sum_{i=1}^n q_i + \bar{q}_0\right) - c(\bar{q}_0) - D$$
$$\geqslant q_0\phi\left(\sum_{i=1}^n q_i + q_0\right) - c(q_0) - D \qquad \text{for all} \quad 0 \leqslant q_0 \leqslant Q. \tag{22}$$

The value of $\bar{q}_0\phi(\sum_{i=1}^n q_i + \bar{q}_0) - c(\bar{q}_0) - D$ determines the nature of the set $B_0(q_{-0})$.

Case 1. $\bar{q}_0\phi(\sum_{i=1}^n q_i + \bar{q}_0) - c(\bar{q}_0) - D < 0$.

In this case, the optimum response for the potential entrant is attained by setting $\delta^* = 0$, $0 \leqslant q_0^* \leqslant Q$. That is $B_0(q_{-0}) = \{(\delta^*, q_0^*) \mid \delta^* = 0, 0 \leqslant q_0^* \leqslant Q\}$.

Case 2. $\bar{q}_0\phi(\sum_{i=1}^n q_i + \bar{q}_0) - c(\bar{q}_0) - D > 0$.

In this case, the optimal response for the potential entrant is attained on setting $\delta^* = 1$, $q_0^* = \bar{q}_0$. That is, $B_0(q_{-0}) = \{(\delta^*, q_0^*) \mid \delta^* = 1, q_0^* = \bar{q}_0\}$.

Case 3. $\bar{q}_0\phi(\sum_{i=1}^n q_i + \bar{q}_0) - c(\bar{q}_0) - D = 0$.

Here the optimal response for the potential entrant is attained by either setting $\delta^* = 0$, $0 \leqslant q_0^* \leqslant Q$ or by setting $0 \leqslant \delta^* \leqslant 1$, $q_0^* = \bar{q}_0$. That is, $B_0(q_{-0}) = \{\delta^* = 0, 0 \leqslant q_0^* \leqslant Q\} \cup \{0 \leqslant \delta^* \leqslant 1, q_0^* = \bar{q}_0\}$.

Clearly for fixed $q_1, \ldots, q_n$ the set $B_0(q_{-0})$ may fall into any of the three cases identified above.

Figure 1 is a sketch of the possible shapes of $B_0(q_{-0})$ corresponding to these three cases.

In Cases 1 and 3, the optimal response sets $B_0(q_{-0})$ are convex and so contractible. And in Case 2, the inverted T-shaped set in Fig. 1 is also contractible. To see this, define the map $h(t, q) : I \times B_0(q_{-0}) \to B_0(q_{-0})$ by

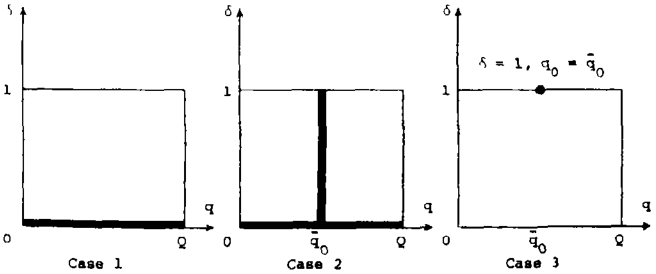


FIG. 1. Optimal response sets for the entrant.

Then $\delta^* = 0$, otherwise $\pi_0^{\delta *}$ would not be maximal for the entrant. But for this situation, the maximization problem reduces to

$$\pi_j^{\delta *} = \max_{0 \leqslant q_j \leqslant Q} \left[q_j \phi \left(\sum_{i \neq j}^{n} q_i^* + q_j \right) - c(q_j) \right], \qquad j = 1,\ldots, n, \tag{32}$$

which is just the maximization problem for the n-firm Cournot oligopoly. Therefore, the output of the established firms is the unique output, ξ_n, of the n-firm oligopoly problem where ξ_n is defined as in (7) and (8). Hence $\delta^* = 0$, $q_1^* = q_2^* = \cdots = q_n^* = \xi_n$ form a noncooperative equilibrium solution to the problem so far as

$$q_0^* \phi(n\xi_n + q_0^*) - c(q_0^*) < D. \tag{33}$$

In particular, the result still holds if

$$D > \max_{0 \leqslant q_0 \leqslant Q} |q_0 \phi(n\xi_n + q_0) - c(q_0)|.$$

And it is straightforward to show that there exists a unique q_0^+, with $0 < q_0^+ < Q - n\xi_n$, which attains the maximum in (33).

And so, for all

$$D > D_2 = q_0^+ \phi(n\xi_n + q_0^+) - c(q_0^+),$$

there exists a noncooperative equilibrium solution to the entry problem with

$$\delta^* = 0, q_1^* = q_2^* = \cdots = q_n^* = \xi_n.$$

Finally, if the equilibrium δ^*, $q_0^*, q_1^*, q_2^*, \ldots, q_n^*$ has $0 < \delta^* < 1$, then

$$q_0^* \phi \left(\sum_{i=1}^{n} q_i^* + q_0^* \right) - c(q_0^*) = D,$$

otherwise, the potential entrant can increase (decrease) δ^* towards 1(0) to improve its profit.

Hence for $0 < \delta^* < 1$, the differentiability conditions for maximum imply the folllowing set of equations:

$$(1 - \delta^*) \left[q_j^* \phi' \left(\sum_{i=1}^{n} q_i^* \right) - c'(q_j^*) + \phi \left(\sum_{i=1}^{n} q_i^* \right) \right]$$
$$+ \delta^* \left[q_j^* \phi' \left(\sum_{i=1}^{n} q_i^* + q_0^* \right) + \phi \left(\sum_{i=1}^{n} q_i^* + q_0^* \right) - c'(q_j^*) \right] = 0,$$
$$j = 1,\ldots, n, \tag{34a}$$

$$q_0^* \phi'\left(\sum_{l=1}^{n} q_l^* + q_0^*\right) + \phi\left(\sum_{l=1}^{n} q_l + q_0^*\right) - c'(q_0^*) = 0 \tag{34b}$$

$$q_0^* \phi\left(\sum_{l=1}^{n} q_l^* + q_0^*\right) - c(q_0^*) = D. \tag{34c}$$

From the symmetry in the problem, the conditions for equilibrium may further be simplified to

$$\delta^* = \frac{q_a \phi'(nq_a) + \phi(nq_a) - c(q_a)}{q_a[\phi(nq_a) - \phi'(nq_a + q_0^*)] + [\phi(nq_a) - \phi(nq_a + q_0^*)]}, \tag{35a}$$

$$q_0^* \phi'(nq_a + q_0^*) + \phi(nq_a + q_0^*) - c'(q_0^*) = 0, \tag{35b}$$

$$q_0^* \phi(nq_a + q_0^*) - c(q_0^*) = D, \tag{35c}$$

where $q_1^* = q_2^* = \cdots = q_n^* = a_a$.

Note that the denominator in (35a) is positive because $\phi(\cdot)$ is nondecreaing and $\phi''(\cdot) \leqslant 0$. Also, from the assumed range of δ^* together with the properties of $H_n(\cdot)$ and other assumptions of the model, it is shown in [10] that

$$q_0^+ < q_0^* < \xi_{n+1} < q_a < \xi_n$$

and that

$$D_2 < D < D_1, \qquad \text{if} \quad 0 < \delta^* < 1.$$

The results on the structure of the noncooperative equilibrium solutions are summarized into a theorem.

THEOREM 7. *Let ξ_n and ξ_{n+1} be the noncooperative equilibrium solutions to the n firm and $n+1$ firm Cournot oligopoly problem with no potential entrants. Also define D_1, D_2 and q_0^+ by*

$$D_1 = \pi^{n+1} = \xi_{n+1}\phi(n\xi_{n+1} + \xi_{n+1}) - c(\xi_{n+1}), \tag{36}$$

$$D_2 = q_0^+ \phi(n\xi_n + q_0^+) - c(q_0^+) = \max_{0<q<Q} [q\phi(n\xi_n + q) - c(q)]. \tag{37}$$

Then the noncooperative equilibrium solution $(\delta^*, q_0^*, q_1^*, \ldots, q_n^*)$ for the Cournot oligopoly probem with one potential entrant may be characterized with respect to entry costs, D, as follows:

Case 1. If $D < D_1$, then $\delta^* = 1$, $q_0^* = q_1^* = \cdots = q_n^* = \xi_{n+1}$ is a noncooperative equilibrium solution.

Case 2. If $D > D_2$, then $-\delta^* = 0$, $q_1^* = q_2^* = \cdots = q_n^* = \xi_n$ is a noncooperative equilibrium solution.

Case 3. If $D_2 < D < D_1$, then in addition to the pure oligopoly solutions identified above there also exists a randomized ebtry solution with $0 < \delta^* < 1$, $q_1^* = q_2^* = \cdots = q_n^* = q_a$ and q_0^* satisfying the set of equations

$$\delta^* = \frac{q_a\phi'(nq_a) + \phi(nq_a) - c'(q_a)}{q_a[\phi'(nq_a) - \phi'(nq_a + q_0^*)] + [\phi(nq_a) - \phi(nq_a + q_0^*)]}, \tag{38a}$$

$$q_0^*\phi'(nq_a + q_0^*) + \phi(nq_a + q_0^*) - c'(q_0^*) = 0, \tag{38b}$$

$$q_0^*\phi(nq_a + q_0^*) - c(q_0^*) = D, \tag{38c}$$

with $q_0' < q_0^* < \xi_{n+1} < q_a < \xi_n$.

Figure 2 is a sketch of the equilibrium solutions characterized in Theorem 7.

The oligopoly problem with entry has a three band equilibrium solution. For entry costs less than D_2, the $(n + 1)$-firm Cournot output solves the problem while for entry costs greater than D_1, the n-firm Cournot output solves it. In other words, if entry costs are moderately low there are no real impediments to entry but once entry costs exceed D_1, entry is forestalled. For intermediate entry costs, a randomized entry solution or any of the pure oligopoly outcomes may be realized.

The pure oligopoly solutions in the middle band are enforceable as threat strategies, one by the potential entrant which announces its intention to enter at the $(n + 1)$-firm oligopoly output, and the other is enforceable by any of the established firms which makes known its intention to produces at the n-firm oligopoly output. The randomized entry equilibrium captures the effects of uncertain entry on the output decisions of established firms. They are forced to reduce their outputs somewhat below what would otherwise be optimal for n noncooperative oligopolists in competition. But the restriction of output serves a dual purpose. It acts as a hedge against overproduction if entry should occur. Yet the restricted output is large enough to discourage entry at the $(n + 1)$-firm oligopoly output.

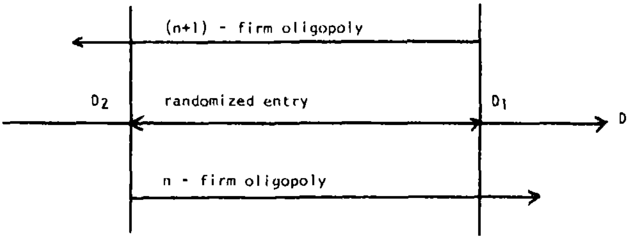


FIG. 2. Variation of equilibrium solution with entry cost.

In the middle band we have three noncooperative solutions whose characteristics can be described qualitatively in terms of threats implicit in the behavior. No entry occurs when the entrant believes that the firms in the industry will stand and fight; entry occurs when those in, believe that new entry is a foregone conclusion and they must adjust to it, and the third case is where all are uncertain.

6. COURNOT OLIGOPOLY WITH FINITE POTENTIAL ENTRANTS

For the model under consideration assume there are n established firms and k potential entrants. Established firms are labelled by the index 1 through n and potential entrants by the index $n+1$ through $n+k$. The variable cost functions $c_j(\cdot)$, $j=1,\dots,n+k$ individually satisfy Assumption 1. In addition, potential entrants face different entry costs. That is, the entry cost of the jth potential entrant is D_{n+j}. Then we have the following general existence theorem.

THEOREM 8. *Under the assumptions of the model, the general Cournot oligopoly problem with n established firms and k potential entrants facing asymmetric entry costs has a noncooperative equilibrium solution of the form* $(\delta_1^*,\dots,\delta_k^*,\ q_1^*,\dots,q_j^*,\dots,q_n^*,q_{n+1}^*,\dots,q_{n+k}^*)$, *where* $0\leqslant\delta_j^*\leqslant 1$, $j=1,\dots,k$, *and* $0\leqslant q_i^*\leqslant Q$, $i=1,\dots,n+k$. q_i^* *is the output of the ith established firm and* δ_j^*, q_{n+j}^*, *respectively, are the entry probability and output of the jth potential entrant.*

Proof. Due to combinatorial reasons, the notation in the proof of the general problem with k potential entrants could get unnecessarily cumbersome. Consequently, we illustrate the proof with only two potential entrants.

Suppose the ith established firm produces q_i, and the jth potential entrant enters with probability δ_j and produces q_{n+j} if it enters; then the expected profits are given by

$$\begin{aligned}\pi_i = {} & (1-\delta_1)(1-\delta_2)\left[q_i\phi\left(\sum_{j=1}^{n} q_j\right)-c_i(q_i)\right]\\ & +\delta_1(1-\delta_2)\left[q_i\phi\left(\sum_{j=1}^{n+1} q_i\right)-c_i(q_i)\right]\\ & +\delta_2(1-\delta_1)\left[q_i\phi\left(\sum_{j=1}^{n} q_j+q_{n+2}\right)-c_i(q_i)\right]\\ & +\delta_1\delta_2\left[q_i\phi\left(\sum_{j=1}^{n+2} q_j\right)-c(q_i)\right], \qquad i=1,\dots,n, \qquad (39a)\end{aligned}$$

$$\pi_{n+1} = \delta_1(1-\delta_2)\left[q_i\phi\left(\sum_{j=1}^{n} q_j + q_{n+1}\right) - c_{n+1}(q_{n+1}) - D_{n+1}\right]$$

$$+ \delta_1\delta_2\left[q_{n+1}\phi\left(\sum_{j=1}^{n} q_j + q_{n+1} + q_{n+2}\right) - c_{n+1}(q_{n+1}) - D_{n+1}\right], \quad (39b$$

$$\pi_{n+2} = \delta_2(1-\delta_1)\left[q_{n+2}\phi\left(\sum_{j=1}^{n} q_j + q_{n+2}\right) - c_{n+2}(q_{n+2}) - D_{n+2}\right]$$

$$+ \delta_1\delta_2\left[q_{n+2}\phi\left(\sum_{j=1}^{n} q_j + q_{n+1} + q_{n+2}\right) - c_{n+2}(q_{n+2}) - D_{n+2}\right]. \quad (39c)$$

The expression for π_i, $i = 1,\dots, n$, π_{n+1} and π_{n+2} are all continuous functions of $q_1,\dots, q_n, \delta_1, \delta_2, q_{n+1}, q_{n+2}$. Also for fixed $\delta_1, \delta_2, q_1,\dots, q_{i-1}, q_{i+1},\dots, q_n, q_{n+1}, q_{n+2}$, π_i is a concave function of q_i, $i = 1,\dots, n$. This follows because π_i is a convex combination of four strictly concave functions of q_i for fixed $\delta_1, \delta_2, q_1,\dots, q_{i-1}, q_{i+1},\dots, q_n, q_{n+1}, q_{n+2}$. Similarly the expressions for π_{n+j} are separately concave in δ_j and q_{n+j} if all other variables other than δ_j and q_{n+j} are fixed, $j = 1, 2$. To see this just rewrite (39b), for example, as

$$\pi_{n+1} = \delta_1\left\{(1-\delta_2)\left[q_{n+1}\phi\left(\sum_{j=1}^{n} q_j + q_{n+1}\right) - c_{n+1}(q_{n+1}) - D_{n+1}\right]\right.$$

$$\left. + \delta_2\left[q_{n+1}\phi\left(\sum_{j=1}^{n} q_j + q_{n+1} + q_{n+2}\right) - c_{n+1}(q_{n+1}) - D_{n+1}\right]\right\}$$

and observe that the expression in the braces is a convex combination of concave functions of q_{n+1} for fixed $\delta_2, q_1,\dots, q_n, q_{n+2}$, so the expression in braces is concave in q_{n+1} if these variables are fixed. Now to prove the theorem, just observe that the same technique used in the proof of Theorem 6 is applicable. Q.E.D.

In general the solution identified in Theorem 8 depends on $D_{n+1},\dots, D_{n+k}$ and all the other data of the problem such as $\phi(\cdot)$, n, k, $c_j(\cdot)$ and Q. The equilibrium solutions also partition entry costs into several bands, but some bands may coalesce depending on the data of the problem.

7. CONCLUSION

We have presented a model of oligopoly which explicitly considers entry barriers and strategic behavior by potential entrants and established firms. Two main results were obtained.

First, we established the existence of a noncooperative equilibrium. At uilibrium, no active firm has an incentive to change production or exit om competition and no firm outside the industry has an incentive to enter. The equilibrium consists of pure strategy outputs for all firms with a ossibility that potential entrants may randomize their entry decisions. Our *existence proof relied on the Eilenberg–Montgomery fixed point theorem rather than that of Kakutani because optimal response sets are not necessarily convex if potential entrants face non-zero entry costs.*

Second, we characterized the equilibrium solutions for a symmetric n-firm oligopoly with one potential entrant and related the equilibrium outcomes to levels of entry costs. Entry costs were endogenously partitioned into three regions: low, moderate and high. When entry costs are low, all firms becomes active as $n + 1$ Cournot oligopolists and when entry costs are high, the potential entrant is excluded, leaving the established firms to operate as n Cournot oligopolists. Uncertainties and subjective factors appear to play a critical role when entry costs are moderate. In this regime, the solutions capture some subtleties in perceptions and behavior. One solution corresponds to a situation where all firms are unsure of each others' intentions and all are cautious; entry is randomized and established firms reduce their outputs. However, if established firms can sufficiently communicate their unwillingness to accommodate entry, an n-firm Cournot olution is enforcible. Likewise, a potential entrant may be accommodated if t can convince the other firms of its intention to penetrate the industry. Thus, the use of threats and other exclusionary tactics become extremely mportant when entry costs are moderate. The more general problem with inite potential entrants may be analyzed and interpreted in a similar way.

A benefit of the approach used in this paper is that it permits an analysis of strategic behavior in the face of entry barriers. Further insights into the ntry problem may be forthcoming if other barriers to entry, such as scale conomies and product differentiation, are studied from a game theoretic erspective.

ACKNOWLEDGMENTS

We are indebted to Matthew J. Sobel for several suggestions and to Pradeep Dubey for a ketch of the mathematical approach to entry used here.

REFERENCES

1. E. Burger, "Introduction to the Theory of Games," Prentice-Hall, Englewood Cliffs, N. J., 1963.
2. P. Dubey and M. Shubik Entry and exit in a closed economic model, unpublished manuscript available from authors, 1978.

3. S. Eilenberg and D. Montgomery, *Fixed point theorems for multivalued transformations*, *Amer. J. Math.* **68** (1945), 214–222.
4. J. W. Friedman, "Oligopoly and the Theory of Games," North-Holland, Amsterdam, 1977.
5. J. J. Gabszewicz and J. P. Vial, Oligopoly "a la Cournot" in a general equilibrium analysis, *J. Econ. Theory* **4** (1972), 381–400.
6. D. W. Gaskins, Jr., Dynamic limit pricing: Optimal pricing and uncertain entry, *J. Econ. Theory* **3** (1971), 306–322.
7. M. I. Kamien and N. L. Schwartz, Limit pricing and uncertain entry, *Econometrica* **39** (1971), 441–454.
8. H. Nikaido and K. Isoda, Notes on noncooperative convex games, *Pacific J. Math.* **5** (1955), 807 – 815.
9. W. Novshek and H. Sonnenschein, Cournot and Walras equilibrium, *J. Econ. Theory* **19** (1978), 223–266.
10. K. O. Nti, "Oligopoly with Potential Entrants," Ph. D. Dissertation, Yale University, December, 1977.
11. K. O. Nti and M. Shubik, "Entry in Oligopoly Theory: A Survey," S. O. M. Working Paper No. 17, Series A, Yale University, 1979.
12. J. Roberts and H. Sonnenschein, On the existence of Cournot equilibrium without concave profit functions, *J. Econ. Theory* **13** (1976), 112–117.
13. M. Shubik, "Strategy and Market Structure," Wiley, New York, 1959.
14. A. M. Spence, Entry, capacity, investment and oligopolistic pricing, *Bell. J. Econ.* **8** (1977), 534–544.

OURNAL OF ECONOMIC THEORY 24, 205–217 1981)

Induced Preferences and the Theory of the Consumer

FRANK MILNE

Department of Economics, Faculty of Economics,
Australian National University, Canberra A.C.T. 2600, Australia

Received November 13, 1979; revised February 20, 1980

A number of authors, e.g., Becker (1965), Stigler and Becker (1977), Lancaster (1966) and Muth (1966) have investigate a generalized theory of the consumer, where the consumer has preferences over characteristics (or commodity objects of choice) which are produced by a wholly owned production technology, using market-traded goods as inputs. This general model can be applied to many different aspects of consumer behaviour (e.g., the consumer allocation of time, the choice of automobile characteristics). Unfortunately it does not appear to be widely recognized, that the Lancaster–Becker theory is a special case of a more general theory of induced preferences. The general theory has been developed in the important contribution of Rader (1964, 1972, 1978), where conditions are provided for the continuity and convexity of induced preferences; and agents with induced preferences are embedded into a general equilibrium system.

In this paper I will draw upon Rader's results, adding some other results on the structure of induced preferences. Section 1 of the paper sets out the basic general theory and the main theorems. We show that given appropriate continuity and convexity assumptions on the underlying utility function and production technology, then we can obtain "neoclassical" induced preferences over the traded goods. Section 2 attempts to show the power of these results in a variety of theoretical applications. Some of these applications are fairly obvious and well known, and I have not considered them in any detail. But there are other applications that may be unfamiliar to the reader, and I have examined three of these in a little more detail. In particular, I have considered induced preferences in (i) Hick's composite good theorem; (ii) financial market theory; and (iii) optimal growth theory.

1. *The General Model*

1.1. Consider a consumer who has a non-empty, attainable consumption set $X \subset A$, where A is a topological space. Let there be a set of traded goods $Z \subset \mathscr{R}^n$ available to the consumer; and a technology correspondence H: $Z \twoheadrightarrow X$.

0022-0531/81/020205-13$02.00/0

Copyright © 1981 by Academic Press, Inc

Assume that the consumer has a continuous utility function defined over his attainable consumption set[1]: that is

A.1. There is a continuous function $u: X \to \mathscr{R}$.

To derive induced preferences on the goods-set Z, we will require some assumptions on the correspondence H.

A.2. For any $z \in Z$, $H(z) \neq \varnothing$;
A.3. H is a continuous correspondence.[2]
A.3′. H is an upper hemi-continuous correspondence.

Now with a direct application of Berge's (1963, Theorem 2, p. 116) we have:

LEMMA 1. *Given* A.1, A.2, A.3′, *then the function* $u^*(z) := \max\{u(x) \mid x \in H(z)\}$ *exists and is upper semi-continuous on* Z.[3]

If we require H to be lower hemi-continuous as well as upper hemi-continuous, then H is continuous, and we have:

LEMMA 2. *Given* A.1, A.2, A.3 *then the function* $u^*(z) := \max\{u(x) \mid x \in H(z)\}$ *exists and is continuous on* Z.

Proof. The Maximum Theorem: Berge (1963, p. 116).

Comment. As Rader (1964) has demonstrated with examples, lower hemi-continuity of H can be destroyed when (a) the underlying production technology is non-convex, or (b) z is an element of the boundary of Z. Thus, in applications, continuity of H is a non-trivial assumption.

Finally, we have the well-known generalization of the Weierstrass theorem:

LEMMA 3. *If* $\hat{Z} \subset Z$ *is compact and* $u^*: Z \to \mathscr{R}$ *is upper semi-continuous, then* u^* *achieves its maximum on* $\hat{Z}$.

Proof. Berge (1963, Theorem 2, p. 76). This concludes our discussion of continuity of u^*.

---

[1] Alternatively I could assume that X is a separable topological space; and there is an irreflexive, transitive and complete preference relation P on $X \times X$, such that P has an open graph. This assumption implies A.1 (see Debreu (1964)). For a discussion of weaker assumptions on the preference relation and derived preferences, see Rader (1978).

[2] For a discussion of continuity notions on correspondences see Berge (1963, Chap. 6).

[3] A function $f: X \to \mathscr{R}$ is said to be upper-semi continuous iff for each $a \in \mathscr{R}$ the set $\{x \in X \mid f(x) \geqslant a\}$ is closed.

1.2. Having discussed continuity, we turn next to non-satiety of u^*. It is easy to show by simple examples that non-satiety of u is insufficient to induce non-satiety on u^*.[4] Therefore we will assume directly that the function u and the correspondence H have the properties sufficient to imply

A.4. For any $z' \in Z$, there exists a $z'' \in Z$ such that $u^*(z'') > u^*(z')$.[5]

1.3. Finally, we can show that if u is quasi-concave, and H is convex, then u^* will also be quasi-concave. Formally,

A.5. A is a vector space.

A.6. X, Z are convex.

A.7. $u\colon X \to \mathscr{R}$ is quasi-concave, i.e., $\{x \in X \mid u(x) \geqslant u(x')\}$ is convex.

A.7′. $u\colon X \to \mathscr{R}$ is concave.

A.8. The correspondence $H\colon Z \twoheadrightarrow X$ is convex, i.e., for $z', z'' \in Z$, then $H(tz' + (1-t)z'') \supset tH(z') + (1-t)\,H(z'')$ for $t \in [0, 1]$.

Now we can prove a slightly modified result of Rader's (1978) Theorem 4.

Lemma 4. *Assuming* A.5, A.6, A.7, A.8 *then* $u^*\colon Z \to \mathscr{R}$ *is quasi-concave.*

Proof. Consider z', $z'' \in P^*(z^*) := \{z \in Z \mid u^*(z) \geqslant u^*(z^*)\}$, $\hat{z} := tz' + (1-t)z''$, $t \in [0, 1]$, and any $x^* \in H(z^*)$. Then there are $x' \in H(z')$ and $x'' \in H(z'')$ such that $u(x') \geqslant u(x^*)$, $u(x'') \geqslant u(x^*)$. Because of A.7, $u(tx' + (1-t)x'') \geqslant u(x^*)$. Also $x_t := tx' + (1-t)x'' \in tH(z') + (1-t)$ $H(z'') \subset H(tz' + (1-t)z'')$ by A.8. Thus for each $x^* \in H(z^*)$, there is an $x_t \in H(tz' + (1-t)z'')$, such that $u(x_t) \geqslant u(x^*)$. Thus $u^*(tz' + (1-t)z'') \geqslant u(z^*)$. ∎

Now given A.7′ a similar proof will produce:

Lemma 5. *Assuming* A.5, A.6, A.7′, A.8 *then* $u^*\colon Z \to \mathscr{R}$ *is concave.*

Finally exploiting the convexity of H, the following lemma is sometimes helpful in proving continuity of a correspondence.

Lemma 6 (Rader (1978)). *A convex correspondence* $H\colon \mathscr{R}^m \twoheadrightarrow \mathscr{R}^n$, *is lower hemi-continuous on the interior of its domain.*

1.4. Given the results in the previous sections, it is obvious that we can induce a standard, continuous (quasi-) concave utility function on Z.

---

[4] For the case of net trades see Rader (1964); for the case of induced preferences over securities see Milne (1976).

[5] Actually Rader uses the weaker requirement of local non-satiation.

Therefore, we can use classical comparative static results and reinterpret them as the consumer's optimal reactions in the traded goods space. For example, the familiar income-compensated demand result becomes:

LEMMA 7. *Given u^* is upper semi-continuous and strictly quasi-concave*[6] *on compact Z: then*

(a) $z(p) := \{z \in Z \mid z(p)$ *Minimizes* pz *on* Z *s.t.* $u^*(z) \geqslant \bar{u}\}$ *is a continuous function on* $\Delta := \{p \in \mathscr{R}^n_+ \mid \Sigma p_h = 1\}$; *and*

(b) $(p'_h - p_h)(z_h(p') - z_h(p) \leqslant 0$.

Proof. Result (a) follows from Berge's Maximum Theorem and the *strict* quasi-concavity of u^*.

Result (b) can be deduced as follows: By definition

(i) $p'(z(p') - z(p)) \leqslant 0$;

(ii) $p(z(p) - z(p')) \leqslant 0$.

Thus (i) and (ii) $\Rightarrow (p' - p)(z(p') - z(p)) \leqslant 0$.

Result (b) follows if p' equals p except for the hth component. ∎

Comment. The result implies that the income-compensated demand function has a non-positive slope.

1.5. In the previous section we exploited the quasi-concavity of u^* to obtain comparative static results. But one can obtain further results by imposing other restrictions on the function u and the correspondence H. For example, assume that H is a linear function.

A.9. Given A.5, A.6, then H: $Z \twoheadrightarrow X$ is a linear *function* such that $H(\alpha z' + \beta z'') = \alpha H(z') + \beta H(z'')$.

Our first result exploits this linearity, along with a restriction on the utility function u. It turns out that if u is homothetic, then under linear H, the induced mapping is also homothetic. Formally,

A.10. u: $X \to \mathscr{R}$ is homothetic. That is, $u = f(g(x))$, where (i) f: $\mathscr{R} \to \mathscr{R}$ is strictly increasing; (ii) g: $X \to \mathscr{R}$ is homogeneous of degree one, i.e., $tg(x) = g(tx)$ for all $x \in X$ and $t > 0$.

Now we can prove a simple lemma:

LEMMA 8. *Given* A.9, A.10, *then u^*: $Z \to \mathscr{R}$ is homothetic.*

---

[6] The function u^*: $Z \to \mathbb{R}$ is strictly quasi-concave iff $\{z \in Z \mid u^*(z) \geqslant \alpha\}$ is strictly convex.

Proof.

$$\begin{aligned} u^*(z) &:= u(H(z)) := f(g(H(z))) \\ &= f(h(z)), \qquad \text{where} \quad h(z) := g(H(z)); \end{aligned}$$

and clearly $h(z)$ is homogeneous of degree one. Thus u^* is homothetic. ∎

Comment. It is well known that homothetic neoclassical preferences imply demand functions $z_h = f_h(p) \cdot W$, where $p \in \mathscr{R}^n_+$ are prices for traded goods and $W \in \mathscr{R}_+$ is consumer wealth.

Our second result uses the linearity of H to induce perfect substitutes over the set of traded goods. (This result has been used extensively in Financial Economics, and, implicitly, in much applied microeconomic theory.)

LEMMA 9. *Given* A.1, A.2, A.9 *and a non-trivial kernel* K *for the linear function* $H\colon \mathscr{R}^n \to A$, *then for any* $\bar{u}^*$, $\{z \in Z \mid u^*(z) \geqslant \bar{u}^*\} \supset \bar{K} \cap Z$, *where* $\bar{K} := \bar{z} + K$, *and* $\bar{z} \in Z$ *satisfies* $u^*(\bar{z}) \geqslant \bar{u}^*$.

Proof. The proof follows from elementary properties of linear algebra.

Comment. The lemma says that if a traded goods "portfolio" can be combined to provide the same consumption as another portfolio, then both portfolios are perfect substitutes under the induced perferences. Furthermore, any feasible linear combination of the two portfolios, is also a perfect substitute.

COROLLARY. *Given the hypotheses of Lemma* 9, *and* $Z := H^{-1}(X)$, *then for any* $\bar{u}^*$, $\{z \in Z \mid u^*(z) \geqslant \bar{u}^*\} \supset \bar{K}$, *where* $\bar{K} := \bar{z} + K$, *and* $\bar{z} \in Z$ *satisfies* $u^*(\bar{z}) \geqslant \bar{u}^*$.

Comment. The corollary is appropriate for security market applications where the traded goods are securities and unbounded short-selling is allowed. (See Milne (1976).)

Now if all consumers face the same linear map, $H\colon \mathscr{R}^n \to A$, then it should be clear that for any consumer $i \in I$, $\{z_i \in Z_i \mid u^*(z_i) \geqslant \bar{u}_i^*\} \supset \bar{K}_i \cap Z_i$, where $\bar{K}_i := \bar{Z}_i + K$, and $\bar{z}_i \in Z_i$ satisfies $u_i^*(\bar{z}_i) \geqslant \bar{u}_i^*$. That is, within the bounds of feasibility imposed by Z and X, every consumer will observe the same indifference relationship (i.e., affine linear manifold) between certain portfolios of traded goods.

It is not difficult to exploit this common linearity, in the preferred sets of consumers, to obtain implications about the relative prices of commodity portfolios on the indifference manifold. For example, consider the following argument:

Definition. A consumer equilibrium is where z^* maximizes $u^*(z)$ for $\{z \in Z \mid pz \leqslant W\}$.

Let preferences be strictly convex, i.e.,

A.7″. If $u(x') > u(x'')$ then $u(tx' + (1-t)x'') > u(x'')$, $t \in (0, 1)$.

Now we have the following lemma:

Lemma 10. *Given* A.1, A.2, A.4, A.6, A.7″, A.9, *a non-trivial kernal* K *for* $H: R^n \to A$, *and the consumer is in equilibrium, such that* $z^* \in \operatorname{int} Z$, *then* $p \in \mathscr{R}^n_+$ *satisfies* $pa = 0$ *for* $a \in K$.

Proof. By assumption, continuous u^* exists. By A.4, A.6, A.7″ the budget constraint holds with equality, i.e., $pz^* = W$ (see Debreu, (1959, p. 71)).

Now $u^*(z^*) = u(H(z^*)) = u(H(z^* + a)) = u^*(z^* + a)$ for appropriate $a \in K$ such that $(z^* + a) \in B_\epsilon(z^*)$.

Assume $pa < 0$; then $p(z^* + a) < W$. But this contradicts the optimality of z^* with budget equality, so that $pa \geqslant 0$. A similar argument, assuming $pa > 0$, implies $pa \leqslant 0$; therefore $pa = 0$. ∎

Comment. This lemma demonstrates that for consumers to hold perfect substitute portfolios of trated goods, then the price of those goods must be in the same linear combination, or otherwise arbitrage is possible.

1.6. Finally we observe that one can induce upon u^* sufficient restrictions to apply standard general equilibrium existence and Pareto Optimality theorems (e.g., Debreu (1959)) for a derived preference economy. Such work has already been accomplished by Rader (see also Milne (1980a)), and we will not pursue it further here.

2. *Applications of Induced Preferences*

2.1. In this section we will provide an outline of some of the established applications of induced preferences. In the subsequent section, we will investigate, in a little more detail, some other applications that demonstrate the power of the approach in quite different areas in economic theory.

The generalized consumer theory developed by Lancaster (1966), Becker (1965) and others, is an obvious application of the induced preference idea. Given suitable assumptions on the consumer's preferences over characteristics and his 'home production function," the model conforms to a special case of our analysis in Section 1. Thus it follows, for example, that one can obtain transparent proofs of the non-positive slope of the income-compensated demand function (our Lemma 7). Furthermore, in using the compensated demand function result, we can use knowledge of the function H to derive partial responses in the optimal characteristics. For example, Becker has used a simple time-constraint to deduce responses in the optimal choice of leisure, when there are variations in wage rates.

Formally, the managerial theory of the firm developed by Williamson (1964) and others, has a similar structure to the generalized consumer theory. In this theory a consumer is a sole owner or manager of a firm, and there are non-traded goods and services flowing between himself and his firm. This model can be analysed as an induced preference problem to obtain obvious comparative static results in the traded goods responses of the owner-firm subeconomy.[7]

Another variant of this model, is the analysis of net trades between countries.[8] The reduced form correspondence H contains production sets, endowments, and the possibility of non-traded goods. Given sufficient assumptions, these complications can be suppressed, and the international trade problem can be analyzed *as if* it was a consumer exchange model.[9]

In standard consumer theory (or producer theory) one can analyze the properties of the indirect utility function, i.e., the derived utility function, treating prices and income as the arguments of the function. Much of this theory of "duality" exploits differentiability— a topic we have omitted here.

Our final example, concerns transaction costs and transaction motives for holding money. Consider a consumer who has preferences over commodities such that he is indifferent to the consumption of gold. Also he has available a transaction technology, where gold is an input. It is not difficult to show, exploiting the ideas of Section 1, that one can induce preferences over traded goods such that gold is now desirable.[10]

2.2. Our first systematic application is the well-known "composite good" theorem first proved by Hicks (1946). The theorem states that given (a) a neoclassical utility function (i.e., quasi-concave, continuous, differentiable etc.); and (b) a partition of the set of goods such that the relative prices for each partition are held constant, then one can derive a neoclassical utility function over the expenditures on each of the subsets of goods. The usual proof—following Hicks—uses calculus methods. Our proof will not assume differentiability, and uses the analysis of Section 1.

---

[7] Other aspects, particularly the questions of optimality and the objective function of the firm, are handled more transparently by a variant of the Lindahl economy idea. See Milne (1980b).

[8] Rader's (1964) original contribution was motivated toward this end. He also perceived more general applications of the induced preference idea, that have been exploited subsequently.

[9] International trade theorists will be familiar with the corresponding diagrammatic idea of net trade indifference curves. See Negishi (1972).

[10] For references to this approach see Nagatani (1978).

Formally, consider the problem:

$$\begin{aligned} &\max_{x \in \mathscr{R}^n_+} u(x) \\ &\text{s.t.} \quad \text{(a)} \quad y_k = \sum_{j \in K_l} p_j x_j, \qquad \text{for } \{K_l\}, \text{ a partition of } \{1, \ldots, n\}. \qquad (1) \\ &\qquad\ \ \text{(b)} \quad y_k \geqslant 0 \quad \text{for all } k = 1, \ldots, K. \\ &\qquad\ \ \text{(c)} \quad p_j > 0 \quad \text{for all } j = 1, \ldots, n. \end{aligned}$$

Now assume

B.1. $y = (y_k) \in Y := \mathscr{R}^K_+ \cap C^K$, where C^K is a compact cube in $\mathscr{R}^K$. The set Y can be considered a set of *attainable* expenditures. The following theorem is a straightforward application of the ideas in Section 1.

THEOREM (Composite Good). *Given problem* (1), A.1, A.7[11], *and* B.1 *then there exists a continuous, quasi-concave, utility function* $y^*(y)$ *on* int(Y).

Proof. Define the correspondence H: $Y \twoheadrightarrow \mathscr{R}^n_+$ by the linear map in (a). Given (b) and (c) it follows that $H(Y)$ is compact. Thus, the graph of H': $Y \twoheadrightarrow H(Y)$ is closed, and therefore upper hemi-continuous. Now Y, $H(Y)$ are convex, and H' is a convex correspondence. By Lemma 6, H' is lower hemi-continuous on int(Y). Thus H' is continuous on int(Y). By Lemma 2 there exists the postulated, continuous u^*: int$(Y) \to \mathscr{R}$. Convexity of H' and A.7, by Lemma 4, imply quasi-concavity of u^*. ▮

Comment. One can induce non-satiation upon u^*, by assuming a suitable non-satiation assumption on u. We will not pursue that argument here.

2.3. In finance theory, induced preferences have been applied extensively to obtain particular demand functions, and relationships between asset prices.

To begin, consider a simple portfolio problem under uncertainty. There is one physical commodity (e.g., corn, or "consumption dollars") which is the consumer's object of consumption. The consumer must choose in a competitive market, a portfolio now, which will provide a random payoff tomorrow in terms of the physical commodity. We will assume that the consumer's preferences conform to the von Neumann–Morgenstern axioms, and exhibit risk-aversion.

Formally, consider a probability space $(\Omega, \mathscr{F}, \mathscr{P})$ and a set of bounded asset return random variables R_j: $\Omega \to \mathscr{R}$, $j \in J := \{1, \ldots, n\}$. Given that the consumer holds a portfolio $a \in A \subset \mathscr{R}^n$, A compact, then the portfolio return is the random variable $R_p := \sum_j R_j a_j$. The consumer is restricted by the budget constraint $pa = W$, where $p \in \mathscr{R}^n_+$, and $W > 0$. Consider u: $\mathscr{R}_+ \to \mathscr{R}$, to be bounded, continuous and concave. Then the consumer's problem can be written as

---

[11] Concavity implies continuity on the *interior* of $\mathscr{R}_+$.

$$\operatorname*{Max}_{\{a\}} U := \int_\Omega u(R_p(\omega)) \, \mathscr{P}(d\omega)$$

subject to

$$\begin{aligned} &\text{(a)} \quad R_p(\omega) = \sum_j R_j(\omega) a_j, \\ &\text{(b)} \quad pa = W, \\ &\text{(c)} \quad A \subset \left\{ a \in \mathscr{R}^n \,\middle|\, \sum_j R_j(\omega) a_j \geqslant 0 \text{ a.e.} \right\}. \end{aligned} \tag{2}$$

Comment. The restriction (c) allows short-selling of securities (i.e., negative holdings of securities) so long as the consumer remains solvent in the feasible states.[12]

By simply substituting the constraint (a) into the objective function, we obtain the trivial induced utility function

$$u^*(a) := \int u \left(\sum_j R_j(\omega) a_j \right) \mathscr{P}(d\omega),$$

which inherits the properties of concavity. Problem (2) therefore, has been reduced to a simple, familiar form. It follows, therefore, that we can apply directly Lemmas 7–10 to the portfolio problem. That is, the income compensated dmand function for an asset has a non-positive slope; if the utility function $u(R_p)$ is homothetic, then so is the derived utility function $u^*(a)$[13]; if the set of assets contains subsets with returns that are linearly dependent, then the prices of those assets must satisfy the same linear relationship. (This last assertion is the foundation for the Modigliani–Miller propositions on invariance of the firm value, when the financial structure is changed.)

The two period analysis can be extended to many periods in a straightforward manner, so that the induced utility function is derived from a stochastic optimizing problem over many periods. This model can be generalized further, to incorporate many consumption goods so that the

---

[12] A more reasonable constraint would be that the consumer must remain solvent in the feasible states as viewed by the lender. For a more detailed discussion, see Milne (1980c).

[13] This statement is analyzed in detail in Milne (1979), where it is shown that homotheticity and additivity is equivalent to a small class of explicit utility functions (i.e., power and logarithmic functions).

theory falls into the general framework of multi-period Temporary Equilibrium.[14] *Of course, Lemmas 7–10 can be applied directly to this more* general model in a straightforward fashion.

Finally, in the more subtle mathematical world of continuous-time-stochastic processes Black and Scholes (1973) have demonstrated an arbitrage theorem relating the price of an option to its underlying stock price and the price of a risk-free bond. In essence, the theorem is a sophisticated application of Lemma 10, where the option price is a linear combination of the stock price and the bond price, the weights being expressed in terms of the underlying probability distribution of the stochastic process, and the time to expiry of the option.[15]

2.4. The last example provides a sketch of a simple result that has been applied in optimal growth theory. This result illustrates the power of the induced preference idea in optimal growth theory.

Gale (1967), McKenzie (1976), Benveniste and Scheinkman (1979, 1981) and others have used the idea of the Valuation function (or induced utility function) for an optimal growth program. Consider a discrete time dynamic problem of the following type:[16]

Given the technology set $T \subseteq \mathscr{R}^{2n}$, find the sequence $\{x_t^*\}$ of capital goods $t = 0, 1, \ldots, \infty$ which solves

$$\begin{aligned} &\text{Maximize } \sum_{t=0}^{\infty} u(x_t, x_{t+1}, t) \\ &\text{subject to } (x_t, x_{t+1}) \in T \qquad \text{for all } t, \text{ and } x_0 \text{ fixed.} \end{aligned} \tag{3}$$

Comment. The utility function $u(x_t, x_{t+1}, t)$ can be induced from a more fundamental utility function over consumption, which is derived from a capital-using technology. This argument is a straightforward application of the analysis in Section 1.

Now assume that there exists a solution $\{x^*(x_0, t)\}$ to the problem (3). Furthermore, assume the valuation function, V, is well defined, where

$$V(x_0) := \sum_{t=0}^{\infty} u(x^*(x_0, t), x^*(x_0, t+1), t).$$

---

[14] Induced preferences have been used widely in the established Temporary Equilibrium literature (see Grandmont (1977)).

[15] Notice that in lemma 10, the set X is a subset of a topological vector space. For a detailed discussion of the relevant topology see Harrison and Kreps (1979), where the arbitrage result is discussed at length.

[16] The continuous-time model can be analyzed in a similar fashion to the arguments below.

Consider the following assumptions:

(a) T is convex, and int $T \neq \emptyset$;

(b) For each t, $u(\ldots, t)$: $T \to \mathscr{R}$ is concave and differentiable.

(c) Given an optimal slution $\{x_t^*\}$ to (3), let $V(\ )$ be well defined for $x \in B_\epsilon(x_0)$.

(d) $(x_0, x^*(x_0, 1)) \in$ int T.

Note. Given (a) and (b), then $V(\ )$ is concave by Lemma 5. Benveniste and Scheinkman (1979) have proved the following result.

LEMMA 11. *Given* (a–d), *then V is a differentiable function at x_0 with*

$$V'(x_0) = \frac{\partial u}{\partial x}(x_1, x^*(x_0, 1)).$$

Because V has a gradient vector V' at x_0, and V is concave, then there exists a non-trivial $p_0 \in \mathscr{R}^n$ such that $p_0 = V'(x_0)$ and

$$p_0(x - x_0) \geqslant V(x) - V(x_0) \qquad \text{for all} \quad x \in B_\epsilon(x_0).$$

Clearly, the condition $p_0 = V'(x_0) = (\partial u/\partial x)(x_0, x^*(x_0, 1))$ has the simple interpretation of the shadow price of capital being equated to the marginal contribution of the capital stock to the value of the program.

Now let us assume that the utility function takes the simple discounted form.

(e) $u(x_t, x_{t+1}, t) = \bar{u}(x_t, x_{t+1})\, \delta^t$, $0 < \delta < 1$.

Finally, define

$$V(x^*(x_0, \tau), \tau) = \delta^\tau \sum_{t=\tau}^{\infty} \bar{u}(x^*(x_0, t), x^*(x_0, t+1))\delta^{t-\tau}.$$

If we assume

(f) $x^*(x_0, \tau) \in B_\epsilon(x_0)$, for all $\tau = 1, 2, \ldots, \infty$,

then, clearly $V(x, \tau) = \delta^\tau V(x)$.

Defining $\psi_\tau = p_\tau \cdot \delta^{-\tau}$, where $p_\tau = V'(x^*(x_0, \tau))\delta^\tau$, then

$$\psi_\tau(x - x^*(x_0, \tau)) \geqslant V(x) - V(x^*(x_0, \tau)).$$

Consider $x^*(x_0, \tau')$, $x^*(x_0, \tau'')$, for $\tau' \neq \tau''$, then

$$\psi_{\tau'}(x^*(x_0, \tau'') - x^*(x_0, \tau')) \geqslant V(x^*(x_0, \tau'')) - V(x^*(x_0, \tau'))$$

and

$$\psi_{\tau''}(x^*(x_0,\tau') - x^*(x_0,\tau'')) \geqslant V(x^*(x_0,\tau')) - V(x_0,\tau'')),$$

imply,

$$(\psi_{\tau'} - \psi_{\tau''})(x^*(x_0,\tau') - x^*(x_0,\tau'')) \leqslant 0. \tag{4}$$

Inequality (4) plays an important role in turnpike-type arguments. (See McKenzie, and Long (1979).) In the special case where $n = 1$, then (4) implies that the optimal capital stock and its current value shadow price move in opposite directions.

Conclusion

In this paper I have tried to show the analytic power of the induced preference idea. Essentially it simplifies messy problems, and provides almost transparent proofs of propositions that are obscure in their primal form. I have not attempted to be exhausitve in this exploration; but I have tried to indicate the power of the approach in diverse problems in economic theory.

ACKNOWLEDGMENTS

I am indebted to N. V. Long, Richard Cornes and a referee for comments. I take sole responsibility for any errors.

REFERENCES

BECKER, G. (1965), A theory of the allocation of time, *Econ. J.* **75**, 493–517.

BENVENISTE, L. M. AND SCHEINKMAN, J. A. (1979), On the differentiability of the value function in dynamic models of economics, *Econometrica* **47**, No. 3, May.

BENVENISTE, L. M. AND SCHEINKMAN, J. A. (1981), Duality theory for dynamic optimization models of economics, *J. Econ. Theory*, **24**, in press.

BERGE, C. (1963), "Topological Spaces," Macmillan Co., New York.

BLACK, F. AND SCHOLES, M. (1973), The pricing of options and corporate liabilities, *J. Pol Econ.*, May/June.

DEBREU, G. (1959), "Theory of Value," Yale Univ. Press, New Haven, Conn.

GALE, D. (1967), On optimal development in a multi-sector economy, *Rev. Econ. Studies* **34** January.

GRANDMONT, J. M. (1977), Temporary general equilibrium theory, *Econometrica*, April.

HARRISON, J. M. AND KREPS, D. M. (1979), Martingales and arbitrage in multiperiod securities markets, *J. Econ. Theory* **20**, 381–408.

HICKS, J. (1946), "Value and Capital," Oxford Univ. Press, London/New York.

LANCASTER, K. (1966), A new approach to consumer theory, *J. Pol. Econ.* **74**, 132–157.

LONG, N. V. (1979), Two theorems on generalized diminishing returns and their applications to economic analysis, *Econ. Record*, March.

McKenzie, L. (1976), Turnpike theory, *Econometrica*, September.

Milne, F. (1976), Default risk in a general equilibrium asset economy with incomplete markets, *Internat. Econ. Rev.* **17**, No. 3, October.

Milne, F. (1980a), "Induced Preferences, Competitive Equilibrium and Optimality," mimeo, Australian National University.

Milne, F. (1980b), The objective function of the firm as a collective choice problem, *Public Choice*, in press.

Milne, F. (1980c), Short-selling, default risk, and the existence of equilibrium in a securities model, *Internat. Econ. Rev.*, June.

Milne, F. (1979), Consumer preferences, linear demand functions and aggregation in competitive asset markets, *Rev. Econ. Studies*, July.

Muth, R. (1966), Household production and consumer demand functions, *Econometrica* **34**, 699–708.

Nagatani, K. (1978), "Monetary Theory," North-Holland, Amsterdam.

Negishi, T. (1972), "General Equilibrium Theory and International Trade," North-Holland, Amsterdam.

Rader, T. (1964), Edgeworth exchange and general economic equilibrium, *Yale Econ. Essays* **4**, 133–180.

Rader, T. (1972), "Theory of Microeconomics," Academic Press, New York/London.

Rader, T. (1978), Induced preferences on trades when preferences may be intransitive and incomplete, *Econometrica*, January.

Stigler, G. and Becker, G. (1977), De Gustibus non est disputandum, *Amer. Econ. Rev.*, March.

Williamson, O. (1964), "The Economics of Discretionary Behavior," Prentice–Hall, Englewood Cliffs, N.J.

JOURNAL OF ECONOMIC THEORY 24, 218-225 (1981)

On the Cores of Economies with Indivisible Commodities and a Continuum of Traders

M. ALI KHAN

Department of Political Economy, The Johns Hopkins University,
Baltimore, Maryland 21218

AND

AKIRA YAMAZAKI

Department of Economics, University of Illinois,
Chicago, Illinois 60680

Received November 28, 1979; revised March 5, 1980

1. INTRODUCTION

We are concerned in this paper with the theory of the core of exchange economies with indivisible commodities and an atomless measure space of agents. For economies with a finite number of agents, contributions to this line of research have been made by Henry [4], Shapley, Scarf [8], and others. (See the references given in [8].)

It is well known that in exchange economies with an atomless measure space of agents and perfectly divisible commodities, the core coincides with the set of competitive allocations and is non-empty. These basic results of Aumann [1, 2] do not extend to commodity spaces with indivisibilities. The point may be best understood in terms of an example due to Mas-Colell [7]. Consider an economy which gives weight one to the preference-endowment combination depicted in Fig. 1. The endowment is denoted by e and the preferences are described by the indifference curves given as the intersections of the dotted lines with the consumption set. It can be checked that in this economy the mean excess demand is bounded away from zero and hence, competitive equilibria do not exist. On the other hand, one can also check without difficulty that an allocation that assigns $(4, 0)$ to half of the consumers and $(0, 4)$ to the remaining half is in the core of the economy.

The approach taken by Mas-Colell to cope with these problems is to introduce a restriction on the endowment distribution of the economy so that the aggregate (mean) demand becomes sufficiently regular. This ensures

0022-0531/81/020218-08$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

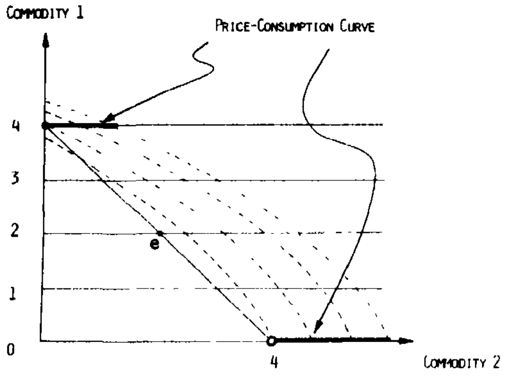


FIGURE 1

existence of an equilibrium as well as the equivalence of the core and the set of competitive allocations. The restriction is that the endowment distribution be "dispersed" in an appropriate sense. (See Mas-Colell [7], and Yamazaki [9, 10].)

The approach taken in this paper is to investigate the core of the economy without any reference to the competitive allocations (but rather to "weakly competitive" allocations). Going back to Mas-Colell's example, note that each agent has a positive amount of the divisible commodity. We shall show in the sequel that this assumption, supplemented by the (overriding) desirability of the divisible commodity, is sufficient for a general theorem on the non-emptiness of the core of an economy. This certainly makes the core a slightly more attractive solution concept than the competitive equilibrium in the presence of indivisible commodities. (Also see Shapley and Scarf [8] on this.)

The paper could be alternatively viewed as an attempt to investigate the extent to which Aumann's basic theorems can be salvaged when we have indivisible commodities and an endowment distribution which is not "dispersed."

The mathematical model of the paper is taken from Mas-Colell [7] which in turn originates in the work of Aumann [1, 2] and Hildenbrand [5]. The proofs of the results are standard in that we simply apply well-known results and techniques found in Hildenbrand [5] to the present model.

The paper is organized as follows. Section 2 is devoted to a precise statement of the results. Theorems 1 and 3 establish the non-emptiness of the core and an approximate core, respectively, under conditions which do not necessarily guarantee existence of a competitive equilibrium. Theorems 2 and 4 are versions of Aumann's basic theorem on the equivalence of the core and the set of competitive allocations. Section 3 is devoted to the proofs of these results and Section 4 ends the paper with a concluding remark.

2. THE MODEL AND RESULTS

Let Z be the set of nonnegative integers and $\Omega = Z^{l-1} \times [0, \infty)$, where l is a positive integer. We shall denote the ith component of an element of Ω by the superscript i and shall use the following convention for the ordering of elements of Ω: $>$, $\geqslant$ and $\geqq$. Let $B_\varepsilon(x)$ be the open ball in Ω centered at x and with radius ε.

A *preference relation* is defined by a pair $(\Omega, \succ)$, where Ω is a *consumption set* and $\succ \subset \Omega \times \Omega$ an irreflexive, transitive, and continuous binary relation defined on Ω. We shall denote the set of all preference relations by $\wp$. The topology on $\wp$ is the one induced by the closed convergence on the closed sets $\Omega \times \Omega \setminus \succ$, $\succ \in \wp$, as in [5, pp. 96–98].

Let $(A, \mathscr{A}, \nu)$ be an atomless, measure space of economic agents with $\nu(A) = 1$. An *economy*, $\mathscr{E}$, is a measurable mapping of the measure space $(A, \mathscr{A}, \nu)$ into the space of agents' characteristics $\wp \times \Omega$ such that $\int_A e\,d\nu$ is finite and strictly positive where e is the projection of $\mathscr{E}(a)$ to Ω and is interpreted as the initial endowment mapping. We shall correspondingly denote the projection of $\mathscr{E}(a)$ on $\wp$ by $\succ_a$ for all a in A.

An *allocation* for $\mathscr{E}$ is an integrable function f from A into Ω such that $\int_A f\,d\nu \leqq \int_A e\,d\nu$. A *coalition* is an element of $\mathscr{A}$ with positive measure. An allocation g is said to *improve upon* an allocation f if there exists a coalition S such that

(a) $g(a) \succ_a f(a)$ for almost all a in S.

(b) $\int_S g\,d\nu \leqq \int_S e\,d\nu$.

The *core* of $\mathscr{E}$, $\mathscr{C}(\mathscr{E})$, is the set of allocations which cannot be improved upon.

For our first result we shall need the following assumptions on $\mathscr{E}$:

A1. *Local Nonsatiation*: For all $\varepsilon > 0$ and for all $x \in \Omega$, there exists $z \in B_\varepsilon(x)$ such that $z \succ_a x$.

A2. *Overriding Desirability of the Divisible Commodity*: For all x, $y \in \Omega$, there exists $z \in \Omega$ such that $z^l > y^l$, $z^i \leqslant y^i$ for all $i \neq l$ and $z \succ_a x$.

A3. For almost all a in A, $e^l(a) > 0$.

Assumption A2 is peculiar to models with indivisible commodities and its role is well understood in the literature. (See Broome [3] and Mas-Colell [7].) We believe an assumption such as A3 is inevitable when the "dispersedness" assumption on the endowment distribution is not made. In any case it is not a strong condition.

We now have

THEOREM 1. *An economy $\mathscr{E}$ satisfying $A1$ to $A3$ has a nonempty core, i.e.,* $\mathscr{C}(\mathscr{E}) \neq \varnothing$.

Our next result shows the equivalence of an approximate core with the set of approximately competitive allocations. Our notion of an approximate core is based on strengthening (b) to $\int_S g\,dv < \int_S e\,dv$. In this case, we shall say that the allocation g *strongly improves* upon the allocation f and shall term the set of allocations which cannot be strongly improved upon as the *weak core* of $\mathscr{E}$, $\mathscr{C}_W(\mathscr{E})$. An allocation f is a *weakly competitive allocation*[1] for $\mathscr{E}$ if there exists an l-vector $p \geqslant 0$ such that for almost all a in A

$$pf(a) \leqslant pe(a) \qquad \text{and} \qquad z \succ_a f(a) \Rightarrow pz \geqslant pe(a). \tag{1}$$

The set of weakly competitive allocations for $\mathscr{E}$ is denoted by $\mathscr{W}_W(\mathscr{E})$. We shall denote by $\mathscr{W}^d_W(\mathscr{E})$ the subset of $\mathscr{W}_W(\mathscr{E})$ for which the price vector p has a positive lth coordinate, i.e., $p^l > 0$. We can now state

THEOREM 2. *For an economy $\mathscr{E}$ satisfying $A1$, the weak core coincides with the set of weakly competitive allocations, i.e.,* $\mathscr{C}_W(\mathscr{E}) = \mathscr{W}_W(\mathscr{E})$.

Given Theorem 1, the non-emptiness of $\mathscr{C}_W(\mathscr{E})$ follows trivially from the fact that $\mathscr{C}_W(\mathscr{E}) \supset \mathscr{C}(\mathscr{E})$. However, in view of condition A1 being the only requirement for Theorem 2, a result on the non-emptiness of $\mathscr{C}_W(\mathscr{E})$ without any of the conditions A1 to A3 is of interest.

THEOREM 3. *An economy $\mathscr{E}$ has a non-empty weak core, i.e.,* $\mathscr{C}_W(\mathscr{E}) \neq \varnothing$.

We may note that the concept of weakly competitive allocations is well known and that it was also discussed previously in the context of large finite economies with non-convex preferences but convex consumption sets; see Hildenbrand [2, p. 202], or Khan [3, p. 550]. The notion of weak core can be found in Hildenbrand [5, Problem 1, p. 140]. Given that the equivalence theorem is valid for atomless economies with convex consumption sets, the need for formulating approximate cores and approximate price equilibria for atomless economies was never felt.

---

[1] Weakly competitive allocations are sometimes also referred to as *pseudo-competitive* allocations.

Theorem 2 responds to the non-equivalence of the set of core and competitive allocations by enlarging *both* notions. If conditions A1 to A3 are satisfied, it is not necessary, however, to enlarge the core to obtain a version of the equivalence theorem. Indeed we have

THEOREM 4. *For an economy* $\mathscr{E}$ *satisfying* $A1$ *to* $A3$, $\mathscr{C}(\mathscr{E}) = \mathscr{W}_w^d(\mathscr{E})$.

We now have two sets of answers to the question as to what can be said for atomless economies with indivisibilities but without "dispersed" endowments. We have Theorems 1 and 4 for settings where A1 to A3 hold and Theorems 2 and 3 for those where only A1 holds. It is also worth pointing out that the assumptions of transitivity and continuity of preferences are not needed for the equivalence results given as Theorems 2 and 4. These assumptions are used in the definition of our economy $\mathscr{E}$.

3. PROOFS

We begin with Theorem 2. The fact that $\mathscr{W}_w(\mathscr{E}) \subset \mathscr{C}_w(\mathscr{E})$ is straightforwardly proved; the converse is a routine modification of the proof of Theorem 1 in Hildenbrand [5, pp. 135–35].[2] This is also true for the proof of $\mathscr{C}_w(\mathscr{E}) \subset \mathscr{W}_w^d(\mathscr{E})$ in the context of Theorem 4. This goes beyond the proof of $\mathscr{C}_w(\mathscr{E}) \subset \mathscr{W}_w(\mathscr{E})$ only in demonstrating the positivity of the price of the divisible commodity, i.e., in establishing $p^l > 0$, which easily follows from the application of assumption A2 on the "overriding desirability of the divisible commodity." Given that $\mathscr{C}(\mathscr{E}) \subset \mathscr{C}_w(\mathscr{E})$, the following proposition completes the proof of Theorem 4.

PROPOSITION 1. *Under* $A3$, $\mathscr{W}_w^d(\mathscr{E}) \subset \mathscr{C}(\mathscr{E})$.

Given Theorems 2 and 4, Theorems 1 and 3 are a consequence of the following result:

PROPOSITION 2. (a) $\mathscr{W}_w(\mathscr{E}) \neq \varnothing$. (b) *Under* $A2$, $\mathscr{W}_w^d(\mathscr{E}) \neq \varnothing$.

We take each of the above propositions in turn.

Proof of Proposition 1. Let $f \in \mathscr{W}_w^d(\mathscr{E})$ and p be an associated price vector. Suppose $f \notin \mathscr{C}(\mathscr{E})$; then there exists an allocation g and a coalition S such that

---

[2] *It is appropriate to mention here a contribution of* Hildenbrand (The core of an economy with a measure space of economic agents, *Rev. Econ. Studies* 35 (1968), 443–452). It was proved there that under A1, $\mathscr{C}(\mathscr{E}) \subset \mathscr{W}_w(\mathscr{E})$ without any convexity assumption on the consumption set.

(i) $g(a) \succ_a f(a)$ for almost all a in S.

(ii) $\int_S g\,d\nu \leqq \int_S e\,d\nu$.

Now consider two possible cases as to whether $\int_S g^l\,d\nu$ is less than or equal to $\int_S e^l\,d\nu$. By definition of $\mathscr{W}^d_w(\mathscr{E})$, (i) implies $p \cdot \int_S g\,d\nu \geqslant p \cdot \int_S e\,d\nu$. Given the positivity of p^l, this leads to a contradiction of our first case. Thus, suppose $\int_S g^l\,d\nu = \int_S e^l\,d\nu$. Assumption A3 guarantees that there exists a set $S' \subset S$, $\nu(S') > 0$ such that $g^l(a) > 0$ for almost all a in S'. Define for all a in S'.

$$Q(a) = \{x \in \Omega \mid x \succ_a f(a)\} \cap \{x \in \Omega \mid x \leqslant g(a), x^l < g^l(a)\}.$$

Given continuity of $\succ_a$ $Q(a)$ is non-empty. Since Q is a measurable correspondence from $(S', \mathscr{A}_{S'}, \nu)$ to Ω, there exists a measurable selection h' such that $h'(a) \in Q(a)$ for almost all a in S'. Let $h(a)$ equal $h'(a)$ for all a in S' and equal to $e(a)$ for all a in $A \setminus S'$. Then h is an allocation which satisfies

(i) $h(a) \succ_a f(a)$ for almost all a in S'.

(ii) $\int_{S'} h\,d\nu \leqslant \int_{S'} e\,d\nu$ with inequality in component l.

We are back in our first case with S' and h substituted for S and g.

Q.E.D.

Proof of Proposition 2. The proof is a modification of a proof presented in Yamazaki [11]. To eliminate the need for frequent cross-referencing, we shall provide a complete argument. We begin with the proof of assertion (a).

Define for each a and $p \in R^l_{++} (\equiv \{p \in R^l \mid p^j > 0 \text{ for each } j\})$,

$$G(a, p) = \{x \in \Omega \mid p \cdot x \leqslant p \cdot e(a) \text{ and } z \succ_a x \Rightarrow p \cdot z \geqslant p \cdot e(a)\}.$$

By a standard argument, $G(a, p) \neq \varnothing$ for $p \in R^l_{++}$ and the correspondence $G(a, \cdot): R^l_{++} \to \Omega$ is upper hemicontinuous for $a \in A$.

For each integer k not smaller than l, let

$$PS_k = \{p \in R^l_{++} \mid p^1 + \cdots + p^l = 1 \text{ and } p^j \geqslant 1/k, j = 1, \ldots, l\}, \tag{1}$$

and define a correspondence F_k of PS_k into R^l by

$$F_k(p) = \int_A \{G(\cdot, p) - e\}\,d\nu \tag{2}$$

for each p in PS_k. It is not difficult to show that $G(\cdot, p)$ has a measurable graph and that it is integrably bounded on PS_k. Thus $F_k(p)$ is non-empty according to [5, Theorem 2, p. 62]. Again by standard arguments (see, e.g., [5, Theorem 3, pp. 62 and 150]), F_k is compact- and convex-valued and upper hemicontinuous for each $k \geqslant l$. Since it is true that $p \cdot z \leqslant 0$ for every z

in $F_k(p)$, we can now apply a fixed point theorem (see e.g., [5, (15), p. 39]) to each F_k. Thus, there exists a pair composed of a price vector and an integrable function of A into Ω, (p_k, f_k), such that

$$f_k(a) \text{ belongs to } G(a, p_k) \qquad \text{for} \quad \text{a.e. } a \text{ in } A. \tag{3}$$

$$p \cdot \int_A (f_k - e)\, dv \leqslant 0 \qquad \text{for every } p \text{ in } PS_k. \tag{4}$$

By taking $p = (1/l, ..., 1/l)$, (4) implies, in particular,

$$\int_A (f_k^1(a) + \cdots + f_k^l(a))\, dv \leqslant \int_A (e^1(a) + \cdots + e^l(a))\, dv. \tag{5}$$

By (5), for all $k \geqslant l$, we have $\int_A f_k \; dv \in \{x \in R^l_+ \mid x^1 + \cdots + x^l \leqslant \int_A (e^1(a) + \cdots + e^l(a))\, dv\}$. The last set is compact. Thus there is a subsequence, still denoted by k, such that $\int_A f_k\, dv \to z^*$ as $k \to \infty$. By Fatou's Lemma in l-dimensions, (see, e.g., [5, Lemma 3, p. 69]), there exists an integrable function f of A into R^l_+ such that (i) $f(a) \in Ls(f_k(a))$, a.e. in A, and (ii) $\int_A f\, dv \leqq z^*$. Let $PS = \{p \in R^l_+ \mid p^1 + \cdots + p^l = 1\}$. Since each p_k belongs to the compact set PS, there is a convergent subsequence of the sequence (p_k), $k \geqslant l$. Without loss of generality one can assume that the sequence itself converges to a vector p in PS. From now on the price vector p is fixed. By (i), one can assume $f_k(a)$ converges to $f(a)$, a.e. in A. Thus $f(a) \in \Omega$ and $p \cdot f(a) \leqslant p \cdot e(a)$, a.e. in A. We now show that a.e. in A, $x \succ_a f(a)$ implies $p \cdot x \geqslant p \cdot e(a)$. Indeed, let $x \in \Omega$ be such that $p \cdot x < p \cdot e(a)$ and $x \succ_a f(a)$. Then, there is a positive integer $k_1 > l$ such that $k > k_1$ implies $p_k \cdot x < p_k \cdot e(a)$. By continuity of $\succ_a$, there is a positive integer $k_2 > k_1$ such that $k > k_2$ implies $x \succ_a f_k(a)$. Hence, for $k > k_2$, we have $p_k \cdot x < p_k \cdot e(a)$ and $x \succ_a f_k(a)$, contradicting the fact that $f_k(a) \in G(a, p_k)$ a.e. in A. Thus, $f(a) \in G(a, p)$ a.e. in A.

Now, (4) implies $p' \cdot z^* \leqslant p' \cdot \int_A e dv$ for every $p' \in \bigcup_k PS_k (\equiv \{p'' \in R^l_{++} \mid (p'')^1 + \cdots + (p'')^l = 1\})$. Thus, one must have $z^* \leqq \int_A e dv$. It then follows from (ii) that $\int_A f dv \leqq \int_A e dv$. Therefore, f is a weak competitive allocation for $\mathscr{E}$, completing the proof of assertion (a)

We turn to the proof of assertion (b). It is enough to show that (p, f) obtained above satisfies, under A2,

$$p^l > 0. \tag{6}$$

Assume $p^l = 0$. Since $\int_A e dv > 0$, there is $S \subset A$ such that $v(S) > 0$ and $p \cdot e(a) > 0$ for $a \in S$. For each a in S, there is a vector y in Ω such that $p \cdot y < p \cdot e(a)$. Thus, by A2, for $a \in S$, there exists a vector z in Ω with the property that $z^l > y^l$, $z^i \leqslant y^i$ for all $i \neq l$, and $z \succ_a f(a)$. As $p \cdot z \leqslant$

$p \cdot y < p \cdot e(a)$, this contradicts the fact that $f(a) \in G(a, p)$, a.e. in A, thus establishing (6). Q.E.D.

4. CONCLUDING REMARK

It is of interest to know whether our theorems on the non-emptiness of core allocations extend to economies with more general non-convex consumption sets as have been studied in [9, 11].

ACKNOWLEDGMENTS

The authors thank Peter Newman, Shin-Ichi Takekuna, and an anonymous referee for their comments.

REFERENCES

1. R. J. AUMANN, Markets with a continuum of traders, *Econometrica* **32** (1964), 39–50.
2. R. J. AUMANN, Existence of competitive equilibria in markets with a continuum of traders, *Econometrica* **34** (1966), 1–17.
3. J BROOME, Existence of equilibrium in economies with indivisible commodities, *J. Econ. Theory* **5**(1972), 224–250.
4. C. HENRY, Market games with indivisible commodities and non-convex preferences, *Rev. Econ. Studies* **39** (1972), 73–76.
5. W. HILDENBRAND, "Core and Equilibria of a Large Economy," Princeton Univ. Press, Princeton, N. J., 1974.
6. M ALI KHAN, Some equivalence theorems, *Rev. Econ. Studies* **61** (1974), 549–565.
7. A. MAS-COLELL, Indivisible commodities and general equilibrium theory, *J. Econ. Theory* **16** (1977), 443–456.
8. L. SHAPLEY AND H. SCARF, On cores and indivisibility, *J. Math. Econ.* **1** (1974), 23–37.
9. A. YAMAZAKI, *An equilibrium existence theorem without convexity assumptions*, *Econometrica* **46** (1978), 541–555.
10. A. YAMAZAKI, On the pseudo-competitive allocations and the core of a large economy, *J. Math. Econ.* **5** (1978), 217–228.
11. A. YAMAZAKI, "Diversified consumption characteristics and conditionally dispersed endowment distribution: regularizing effect of aggregation and existence theorems; mimeo, University of Illinois, Chicago, May 1979.

JOURNAL OF ECONOMIC THEORY 24, 226–239 (1981)

Directional and Local Electoral Equilibria with Probabilistic Voting*

PETER COUGHLIN† AND SHMUEL NITZAN‡

IMSSS, Stanford University, Stanford, California 94305

Received October 17, 1978; revised June 7, 1980

1. INTRODUCTION

Electoral competitions and simple voting games in which each of the candidates and/or the society is restricted to choosing among "local" options (viz. directions of motion away from the status quo or policies within a small neighborhood of the status quo) have been studied by Plott [1967], Kramer and Kleverick [1974], Schofield [1978], Matthews [1979] and Cohen and Matthews [1980]. Such social choice situations are of interest for a number of reasons. For example, the expense of acquiring information may restrict candidates to learning about voter behavior only near the status quo. Alternatively, institutional restrictions may rule out large changes. These and other reasons have been carefully discussed by the above authors.

The existing results on this problem have been derived for societies with full participation electorates and deterministic voting. These results have established that there are usually directional, stationary and local electoral equilibria when the society's policy space is one-dimensional. However, they have also established that such equilibria rarely exist when there are two or more dimensions. Local cycles have therefore been studied for societies with such multi-dimensional policy spaces.

Recent work on the theory of economic policy formation through elections has been concerned with the consequences of random voting behavior. Full participation electorates have been studied in Comaner [1976], Hinich [1977, 1978], Kramer [1978] and Coughlin and Nitzan [1981]. Related

---

[*] This work was supported in part by the National Science Foundation and by Office of Naval Research Grant No. ONR-N00014-79-C-0685. An earlier version was released as Technical Report No. 303 in the Economics Series of the Institute for Mathematical Studies in the Social Sciences at Stanford University.

[†] Current address: Institute of Economics and Statistics, Oxford University, Oxford, England.

[‡] Current address: Department of Economics, Hebrew University, Jerusalem, Israel.

0022-0531/81/020226-14$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

studies of electorates with abstentions are due to Hinich *et al.* [1972, 1973], McKelvey [1975] and Denzau and Kats [1977].

This paper extends the foundations of spatial models of electoral competition with probabilistic voting from those established in Comaner [1976], Hinich [1977], and Kramer [1978] (Section 2).

From these foundations, we first derive (in Section 3) a necessary and sufficient condition for two directions from a given status quo to constitute an equilibrium strategy pair for candidates in an electoral competition (Theorem 1). This condition implies that there is such an equilibrium pair at every status quo (Corollary 1).

We then derive a necessary and sufficient condition for a status quo to have equilibrium directions which are "no change," (i.e., to have a stationary electoral equilibrium) (Theorem 2). This condition implies the existence of a status quo with a stationary electoral equilibrium (Corollary 2).

In Section 4, we are concerned with local electoral equilibria. We analyze this problem under conditions which are standard in micro-economic analyses. For these societies we derive necessary and sufficient conditions for a status quo to have a local electoral equilibrium (Theorem 3). This condition implies the existence of a status quo with a local electoral equilibrium (Corollary 3).

The results derived in this paper hold for any multi-dimensional policy space and do not use any special symmetry requirements on the distribution of voters' preferences. The general existence results established in Corollaries 1-3 are, therefore, in sharp contrast to the conclusion of Plott [1967], *et al.* (for deterministic voting) that directional and stationary electoral equilibria rarely exist when the society's policy space has at least two dimensions. They also do not use any of the special concavity conditions in Hinich *et al.* [1972], [1973] and Denzau and Kats [1977].

Theorems 2 and 3 also establish the equivalence between status quos which have stationary equilibria (respectively, status quos which have local equilibria) and the stationary outcomes (respectively, local maxima) for the society's mean (or social) log-likelihood function. They, therefore, reveal a close connection between certain electoral choices and those from alternative social choice mechanisms which are based on such functions. This also provides a method for finding the status quos which have electoral equilibria.

Section 5 concludes. The proofs are given in the Appendix.

2. PROBABILISTIC VOTING AND ELECTORAL COMPETITIONS

Empirical studies of voting behavior as a function of proposed policies and existing economic conditions leave a significant amount of unexplained variation (e.g., Kramer [1971], Stigler [1973] and Fair [1978]). This has led to the conclusion that uncertainty and non-policy considerations result in random (or indeterminate) voting behavior when this behavior is viewed as a function of existing and proposed policies. Intriligator [1973], Fishburn [1975] and others have formulated indeterminateness in voter behavior with individual choice probabilities. For a non-empty Euclidean space of social alternatives, $X \subset R^n$, these individual choice probabilities can be summarized by a density function on X (e.g., Nitzan [1975]). This probabilistic voting density function for an individual expresses the probabilities of his choosing an alternative in different possible subsets of X, given that he can determine the social choice unilaterally.

Learning the behavior of every individual is impossible. Therefore, political entrepreneurs (or candidates) have to estimate voter behavior. Since the candidates usually have access to the same information (polls, past election data, etc.), we will assume that candidates obtain a common *probabilistic voting estimator* which estimlates the proportions of the population that are described by particular probabilbistic voting density function. To be precise, let $\Theta \subseteq R^l$ denote an index set of parameters for a class of density functions. Then we are assuming that the candidates obtain a *probabilistic voting estimator*, $\hat{g}(\theta)$, which is a (bounded) probability density function on Θ.

For instance, candidates may be willing to use normal probabilistic voting estimators. In this case, the candidates can estimate the proportions of the population that fall into certain combinations of possible means and variances.

$f(x;\theta)$ will denote a real-valued density function on X which has the parameter $\theta \in \Theta$. Since these density functions are only estimates of individual behavior, we will make three regularity assumptions. Specifically,

(i) Each $f(x;\theta)$ is a positive and continuously differentiable function of x.

(ii) $f(x;\theta)$ is a bounded, measurable function of θ whose lower bound is strictly positive.

(iii) $\partial f(x;\theta)/\partial x_h (h = 1,\dots, n)$, is a bounded, measurable function of θ.

In an election, each voter must choose from a pair of policy-proposing candidates. Therefore, we must learn the consequences of probabilistic voting behavior for pairwise choices. We will let ψ_1 and ψ_2 denote policies which have been proposed by candidates who are labelled 1 and 2, respectively. Additionally, we will let $P^j_\theta(\psi_1, \psi_2)$ denote the probability that an individual, whose behavior is described by $f(x;\theta)$, votes for candidate j, $j = 1, 2$.

The following two assumptions provide the basis for calculating individual binary choice probabilities from probabilistic voting behavior on X.

First, we assume that the candidates estimate the behavior of concerned citizens who vote. This is especially plausible in a local equilibrium where only alternatives near the currently proposed positions are of interest. This means that we are concerned with a *full participation electorate* where

$$P_\theta^1(\psi_1, \psi_2) + P_\theta^2(\psi_1, \psi_2) = 1 \tag{1}$$

for every $\theta \in \Theta$ and $(\psi_1, \psi_2) \in X \times X$. This assumption is repeatedly made in the literature on local equilibria.

Second, we want to relate random voting behavior on a Euclidean policy space to behavior on binary choices, i.e., choices between two proposed policies. We therefore assume that each individual's choice probabilities satisfy (in the terminology of Luce) *independence from irrelevant alternatives*, i.e.,

$$\frac{P_\theta^1(\psi_1, \psi_2)}{P_\theta^2(\psi_1, \psi_2)} = \frac{f(\psi_1; \theta)}{f(\psi_2; \theta)} \tag{2}$$

for each $\theta \in \Theta$, $(\psi_1, \psi_2) \in X \times X$.

This is a "separability" condition which merely states that the relative likelihoods of choosing ψ_1 and ψ_2 from X are preserved when choosing from the set $\{\psi_1, \psi_2\}$. This is the continuous version of the independence from irrelevant alternatives which follows from the basic choice axions in Luce [1959, Axiom 1 and Lemma 3] (also in Luce and Suppes [1965] and Ray [1973]).

Finally, since vote totals are random, a candidate could concern himself with his expected plurality or with his probability of winning. However, Hinich [1977, pp. 212–213] has proven that, for many large electorates with a reasonably large amount of indeterminateness, these two objectives are equivalent. Therefore, we will consider the more tractable objective function, and assume that each candidate is concerned with his expected plurality. We will denote the expected plurality for candidate 1 at the pair of proposed policies $(\psi_1, \psi_2) \in X \times X$ by $Pl(\psi_1, \psi_2)$.

$S \subset X$ will be used to denote the set of policies which are feasible for the society when the candidates' strategic decisions are to be made. We will assume throughout that S is non-empty, compact and contained in the interior of X. When candidates have complete mobility within this set of feasible alternatives, electoral competitions are two-person, zero-sum games where the common strategy set for the two candidates is S and $Pl(\psi_1, \psi_2)$ and $-Pl(\psi_1, \psi_2)$ are the payoff functions for candidates 1 and 2, respectively. This paper analyzes the directional and local electoral competitions that arise within this framework when both candidates start at a (common) status quo and then only evaluate their positions on the margin or in comparison with local alternatives. In these games the payoffs will be either expected pluralities very close to the status quo or marginal variations in expected pluralities.

3. DIRECTIONAL AND STATIONARY EQUILIBRIA

A society always has some current state (or status quo). In this section, we analyze situations in which (i) candidates can, at most, marginally vary their positions from the status quo, and (ii) the status quo is associated with a candidate if he does not choose a direction which varies his position away from it.

The possible directions in which a candidate can move his position from the status quo, together with the choice of "no change," define the candidates' strategy space. The marginal gains in expected plurality from simultaneous variations in positions define the candidates' payoffs. The formal definitions are as follows.

A feasible direction u at a status quo, $\psi \in S$, is a vector in R^n of unit length for which there exists a $\lambda_1 > 0$ such that $\psi + \lambda \cdot u \in S$ for every $\lambda \in (0, \lambda_1)$. The common strategy set, $T(\psi)$, at a status quo $\psi \in S$, consists of all the feasible directions from ψ into S together with the zero in R^n (i.e., together with "no change"). $u \in T(\psi)$ and $v \in T(\psi)$ will denote directions selected by candidates 1 and 2, respectively.

The payoff functions at ψ are defined by the net effects of the simultaneous variations u and v. In symbols, the payoff function for candidate 1, $P(u, v)$ on $T(\psi) \times T(\psi)$, is

$$P(u, v) = D_{(u,v)} Pl(\psi_1, \psi_2) = \sum_{h=1}^{n} \frac{\partial Pl}{\partial \psi_{1h}} u_h + \sum_{h=1}^{n} \frac{\partial Pl}{\partial \psi_{2h}} v_h \tag{3}$$

at $\psi_1 = \psi_2 = \psi$. The payoff function for candidate 2 is given by $-P(u, v)$. $(u^*, v^*) \in T(\psi) \times T(\psi)$ is a *directional electoral equilibrium* (in pure strategies) at the status quo $\psi \in X$ if and only if

$$P(u, v^*) \leqslant P(u^*, v^*) \leqslant P(u^*, v) \qquad \forall u, v \in T(\psi). \tag{4}$$

We can now characterize the directional electoral equilibria which occur at a given status quo.

THEOREM 1. *$(w^*, z^*) \in T(\psi) \times T(\psi)$ is a directional electoral equilibrium at the status quo ψ if and only if*

$$D_{(w^*,0)} Pl(\psi_1, \psi_2) = \max_{u \in T(\psi)} D_{(u,0)} Pl(\psi_1, \psi_2),$$

and

$$D_{(0,z^*)} Pl(\psi_1, \psi_2) = \min_{v \in T(\psi)} D_{(0,v)} Pl(\psi_1, \psi_2)$$

at $\psi_1 = \psi_2 = \psi$.

This also implies

COROLLARY 1. *There exists a directional electoral equilibrium at every status quo in S.*

If, in a directional electoral equilibrium, both candidates choose to remain at the status quo, then the society is in a stationary equilibrium. In our terminology, there is a *stationary electoral equilibrium* at a status quo ψ if and only if the zero vector in R^{2n} is a directional electoral equilibrium at ψ.

We should remark that there are many alternatives to the existing institution of political elections for making a social choice. Various suggestions for aggregating individual choice probabilities and then determining the social choices have been proposed (e.g., see Intriligator [1973], Fishburn [1975] and Nitzan [1975]). We present below one such possible aggregation rule which, it turns out, is closely related to electoral competitions.

Recall that individual choice probabilities are summarized in the density functions $f(x;\theta)$, or, equivalently, in the log-likelihood functions $\ln f(x;\theta)$. The society's *mean* (or *social*) *log-likelihood function* is given by

$$L(x)=\int_{\Theta} \ln(f(x;\theta))\,\hat{g}(\theta)\,d\theta. \tag{5}$$

If a society is concerned with the marginal changes in $L(x)$ which result from variations in the status quo, then they will not vary a status quo $\psi \in X$ (i.e., will choose "no change") if and only if

$$D_u L(\psi) \leqslant 0, \qquad \forall u \in T(\psi). \tag{6}$$

We will therefore refer to any such point in the society's policy space as a *stationary outcome* for the social (or mean) log-likelihood function.

Using this alternative social choice mechanism and Theorem 1 we can now characterize the status quos at which a stationary electoral equilibrium occurs.

THEOREM 2. *There is a stationary electoral equilibrium at a status quo $\psi \in X$ if and only if ψ is a stationary outcome for the society's mean log-likelihood function.*

This, in turn, directly implies

COROLLARY 2. *There exists some status quo, $\psi \in X$, at which there is a stationary electoral equilibrium.*

4. Local Equilibria

Some stationary equilibria have the local instability property that, if the initial position of the candidates is perturbed slightly, we can find the candidates choosing an equilibrium direction which points away from the neighborhood stationary equilibrium point. This is the case, for instance, when candidates are locally minimizing their expected pluralities (taking their rival's position as no change from the status quo). We therefore now consider existence and location questions for the stronger concept of a *local Nash equilibrium*.

We can consider the possible candidate strategies to be "small" neighborhoods of the status quo, and the payoffs to be the candidates' respective expected pluralities rather than marginal changes in their values. Given the expected plurality function, $Pl(\psi_1, \psi_2)$, the status quo $\psi \in X$ is a *local electoral equilibrium* if and only if

$$Pl(\psi_1, \psi) \leqslant Pl(\psi, \psi) = 0 \leqslant Pl(\psi, \psi_2) \tag{7}$$

for every $\psi_1 \in N_{\varepsilon_1}(\psi)$ and $\psi_2 \in N_{\varepsilon_2}(\psi)$ for some $\varepsilon_1, \varepsilon_2 > 0$.

We will study local electoral equilibria under assumptions on the "social opportunity set" S which are standard for such sets in microeconomic analysis. Specifically, we will assume that S is a compact subset of X which is defined by twice continuously differentiable equations of the form

$$g_k(x) = 0 \qquad (k = 1,..., m < n) \tag{8}$$

with the property that each $x \in S$ is a normal point (e.g., see Hestenes [1975, p. 114]).

This will mean that candidates (when they take their rival's position as given) and society will have Lagrangean maximization problems.

We will also strengthen our regularity conditions (see Section 2) on the individuals' probabilistic voting estimators and additionally assume:

(1) Each $f(x; \theta)$ is twice continuously differentiable in x, and

(2) $\partial^2 f(x; \theta)/\partial x_h \partial x_l (h, l \in \{1,..., n\})$ is a bounded, measurable function of θ.

Finally, in order to study this problem for C^2 payoff functions, we will analyze those situations in which $Pl(x, \psi)$ and $L(x)$ are nondegenerate functions with respect to the constraints given by (8) (see Hestenes [1975], p. 153). It should be observed that this imposes essentially no further restriction on the class of electoral competitions we are analyzing since, generically, every C^2 function is non-degenerate (e.g., see Hirsch [1976, Theorem 6.1.2]).

For this standard microeconomic decision-making context we can obtain a complete characterization of the status quos at which there are local electoral equilibria:

THEOREM 3. *There is a local electoral equilibrium at a status quo $\psi \in S$ if and only if ψ is a local maximum for the society's mean log-likelihood function.*

This, again, gives us a general equilibrum result:

COROLLARY 3. *There exists some status quo, $\psi \in S$ at which there is a local electoral equilibrium.*

5. CONCLUSION

This paper has extended the foundations of spatial models of electoral competitions with probabilistic voting. In particular, we have included infinite voter populations, estimators of voting behavior for the candidates and multi-dimensional policy spaces. We then derived necessary and sufficient conditions for directional, stationary and local electoral equilibria. These conditions establish general existence results for such electoral equilibria. They also reveal the close relation between elections and certain social choice mechanisms which select stationary outcomes or local maxima of the society's mean log-likelihood function.

APPENDIX

Proof of Theorem 1. First, we obtain the expected plurality function. By (1) and (2),

$$P_\theta^1(\psi_1, \psi_2) = \frac{f(\psi_1; \theta)}{f(\psi_1; \theta) + f(\psi_2; \theta)} \tag{9}$$

for every $\theta \in \Theta$ and $(\psi_1, \psi_2) \in X \times X$.

The expected plurality for candidate 1 from an individual with the probabilistic voting density function indexed by θ is

$$P_\theta(\psi_1, \psi_2) = P_\theta^1(\psi_1, \psi_2) - P_\theta^2(\psi_1, \psi_2) = 2P_\theta^1(\psi_1, \psi_2) - 1.$$

Therefore, the expected plurality for candidate 1 from the entire population is

$$
\begin{aligned}
Pl(\psi_1, \psi_2) &= \int_{\Theta} P_\theta(\psi_1, \psi_2)\, \hat{g}(\theta)\, d\theta \\
&= \int_{\Theta} \left\{ 2 \cdot \frac{f(\psi_1; \theta)}{f(\psi_1; \theta) + f(\psi_2; \theta)} - 1 \right\} \hat{g}(\theta)\, d\theta
\end{aligned}
\tag{10}
$$

Since $f(x; \theta)$ is a measurable function of θ for any given $x \in X$, $f(x; \theta) + f(y; \theta)$ is measurable for any given $x, y \in X$. $f(x; \theta) > 0$ for every $x \in X$ and $\theta \in \Theta$. Therefore,

$$\left\{ \theta \in \Theta \;\middle|\; \frac{1}{f(x; \theta) + f(y; \theta)} < k \right\} = \left\{ \theta \in \Theta \;\middle|\; f(x; \theta) + f(y; \theta) > \frac{1}{k} \right\}$$

for any $k > 0$. Furthermore

$$\left\{ \theta \in \Theta \;\middle|\; \frac{1}{f(x; \theta) + f(y; \theta)} < k \right\} = \varnothing$$

for any $k \leqslant 0$.

Hence, $1/[f(x; \theta) + f(y; \theta)]$ and, in turn, $P_\theta(\psi_1, \psi_2)$ are measurable functions of θ. Additionally, by (9), $P_\theta(\psi_1, \psi_2)$ is bounded by 0 and 1. Finally, since $\hat{g}(\theta)$ is a (bounded) probability density function on Θ, the product $P_\theta(\psi_1, \psi_2) \cdot \hat{g}(\theta)$ is bounded and measurable and hence, Lebesgue integrable. Therefore, the integral $Pl(\psi_1, \psi_2) = \int_\Theta P_\theta(\psi_1, \psi_2)\, \hat{g}(\theta)\, d\theta$ exists.

By Corollary 5.9 in Bartle [1966] and the regularity conditions on the $f(x; \theta)$, the gradient of $Pl(x, \psi)$ has as its terms the values given by

$$\frac{\partial Pl(x, \psi)}{\partial x_h} = \int_\Theta \frac{\partial}{\partial x_h} \left(\frac{2f(x; \theta)}{f(x; \theta) + f(\psi; \theta)} - 1 \right) \hat{g}(\theta)\, d\theta \tag{11}$$

for $h = 1, \ldots, n$.

By (10), $Pl(x, y) = -Pl(y, x)$. Therefore, (3) implies $P(w, z) = -P(w, z)$ and $P(z, z) = 0$ for all $w, z \in T(\psi)$. In other words, the candidates' directional game is symmetric and zero-sum. Consequently, $(w^*, z^*) \in T(\psi) \times T(\psi)$ is a directional electoral equilibrium if and only if both (w^*, w^*) and (z^*, z^*) are also directional electoral equilibria. Additionally, this implies that (w^*, w^*) is a directional electoral equilibrium at ψ if and only if

$$D_{(u, w^*)} Pl(\psi, \psi) \leqslant D_{(w^*, w^*)} Pl(\psi, \psi) = 0 \tag{12}$$

for every $u \in T(\psi)$.

By (3), (12) is equivalent to

$$\sum_{h=1}^{n} \frac{\partial Pl(\psi,\psi)}{\partial \psi_{1h}} u_h + \sum_{h=1}^{n} \frac{\partial Pl(\psi,\psi)}{\partial \psi_{2h}} w_h^* \leqslant 0 \tag{13}$$

for every $u \in T(\psi)$.

Now, $Pl(x, y) = -Pl(y, z)$ implies

$$\frac{\partial Pl(\psi,\psi)}{\partial \psi_{2h}} = \frac{-\partial Pl(\psi,\psi)}{\partial \psi_{1h}}.$$

Therefore, (13) is equivalent to

$$\sum_{h=1}^{n} \frac{\partial Pl(\psi,\psi)}{\partial \psi_{1h}} u_h - \sum_{h=1}^{n} \frac{\partial Pl(\psi,\psi)}{\partial \psi_{1h}} w_h^* \leqslant 0 \tag{14}$$

for every $u \in T(\psi)$.

Finally, (14) is equivalent to

$$D_{(u,0)} Pl(\psi,\psi) \leqslant D_{(w^*,0)} Pl(\psi,\psi) \tag{15}$$

for every $u \in T(\psi)$.

By a similar argument, (z^*, z^*) is a directional electoral equilibrium at ψ if and only if

$$D_{(0,v)} Pl(\psi,\psi) \geqslant D_{(0,z^*)} Pl(\psi,\psi) \tag{16}$$

for every $v \in T(\psi)$.

Hence the theorem follows. Q.E.D.

Proof of Corollary 1. At any $\psi \in X$,

$$D_u Pl(x,\psi)|_{x=\psi} = \sum_{h=1}^{n} \frac{\partial Pl(x,\psi)}{\partial \psi_{1h}}\bigg|_{x=\psi} \cdot u_h$$

is linear as a function of the vectors $u \in R^n$. Additionally, since S is compact, the set of feasible directions, at any $\psi \in S$, is a compact subset of R^n. Therefore, $D_u Pl(\psi, \psi)$ achieves a maximum over this set. Either a maximizing direction in this set or the zero vector must, therefore, maximize $D_u Pl(\psi, \psi)$. Q.E.D.

Proof of Theorem 2. First of all, $\ln f(x;\theta)$ is a bounded function of θ which is defined for every $x \in X$, since $f(x;\theta) > 0$ is a bounded function of θ whose lower bound is strictly positive, for any $x \in X$ (see Section 2). Additionally, for any given x,

$$\{\theta \in \Theta \mid \ln f(x;\theta) > k\} = \{\theta \in \Theta \mid f(x;\theta) > e^k\}$$

for any $k \in R$. Therefore, since $f(x;\theta)$ is a measurable function of θ, $\ln f(x;\theta)$ is measurable. Consequently the integral defining $L(x)$ exists.

By Corollary 5.9 in Bartle [1966] and the regularity conditions on the $f(x;\theta)$, for $h = 1,\ldots, n$,

$$\frac{\partial L(x)}{\partial x_h} = \int_\Theta \frac{\dfrac{\partial f(x;\theta)}{\partial x_h}}{f(x;\theta)} \hat{g}(\theta)\, d\theta. \tag{17}$$

As stated in (11),

$$\frac{\partial Pl(x,\psi)}{\partial x_h} = \int_\Theta \frac{\partial}{\partial x_h}\left(\frac{2f(x;\theta)}{f(x;\theta)+f(\psi;\theta)} - 1\right) \hat{g}(\theta)\, d\theta. \tag{18}$$

Therefore,

$$\begin{aligned} \left.\frac{\partial Pl(x,\psi)}{\partial x_h}\right|_{x=\psi} &= 2\cdot\int_\Theta \left.\frac{f(\psi;\theta)\cdot\dfrac{\partial f(x;\theta)}{\partial x_h}}{(f(x;\theta)+f(\psi;\theta))^2}\right|_{x=\psi} \hat{g}(\theta)\, d\theta \\ &= \int_\Theta \frac{\left.\dfrac{\partial f(x;\theta)}{\partial x_h}\right|_{x=\psi}}{2f(\psi;\theta)} \hat{g}(\theta)\, d\theta. \end{aligned} \tag{19}$$

Hence,

$$\left.\frac{\partial Pl(x,\psi)}{\partial x_h}\right|_{x=\psi} = \frac{1}{2}\left.\frac{\partial L(x)}{\partial x_h}\right|_{x=\psi}. \tag{20}$$

Therefore, for any $u \in T(\psi)$,

$$D_{(u,0)}Pl(\psi,\psi) = \sum_{h=1}^{n} \frac{\partial Pl(\psi,\psi)}{\partial x_h} u_h = \frac{1}{2}\sum_{h=1}^{n} \frac{\partial L(\psi)}{\partial x_h} = \frac{1}{2} D_u L(\psi). \tag{21}$$

By Theorem 1, $(0,0) \in T(\psi) \times T(\psi)$ is an equilibrium strategy pair at ψ if and only if $D_{(u,0)}Pl(\psi,\psi) \leqslant 0$ for every feasible direction at ψ and, hence, if and only if $D_u L(\psi) \leqslant 0$ for every feasible direction at ψ. Q.E.D.

Proof of Corollary 2. Consider any $\psi \in X$. Since $f(x;\theta)$ is a continuous function of x (by regularity condition (i)), $\ln f(x;\theta)$ is a continuous function of x. Therefore, $L(x)$ is continuous (by the Lebesgue Dominated Convergence Theorem). Hence, since S is compact, there is some $\psi \in S$ such that ψ is a critical point or a boundary maximum of $L(x)$. Any such point must satisfy (6). Therefore there is some stationary outcome for the social log-likelihood function. Hence, by Theorem 1, there must be a stationary electoral equilibrium. Q.E.D.

Proof of Theorem 3. First, at $x=\psi$,

$$\frac{\partial^2 Pl(x,\psi)}{\partial x_h\,\partial x_l}=\int_\Theta\left\{\frac{2(f(x;\theta)+f(\psi;\theta))^2 f(\psi;\theta)\dfrac{\partial^2 f(x;\theta)}{\partial x_h\,\partial x_l}}{(f(x;\theta)+f(\psi;\theta))^4}-\frac{4f(\psi;\theta)(f(x;\theta)+f(\psi;\theta))\dfrac{\partial f(x;\theta)}{\partial x_h}\cdot\dfrac{\partial f(x;\theta)}{\partial x_l}}{(f(x;\theta)+f(\psi;\theta))^4}\right\}\hat{g}(\theta)\,d\theta \tag{22}$$

(by Corollary 5.9 in Bartle [1966] and the regularity conditions on the $f(x;\theta)$). Therefore,

$$\left.\frac{\partial^2 Pl(x,\psi)}{\partial x_h\,\partial x_l}\right|_{x=\psi}=\int_\Theta\left\{\frac{\dfrac{\partial^2 f(x;\theta)}{\partial x_h\,\partial x_l}f(x;\theta)-\dfrac{\partial f(x;\theta)}{\partial x_h}\cdot\dfrac{\partial f(x;\theta)}{\partial x_l}}{2f(x;\theta)^2}\right\}_{x=\psi}\hat{g}(\theta)\,d\theta. \tag{23}$$

But, at $x\in\psi$ we also have

$$\frac{\partial^2 L(x)}{\partial x_h\,\partial x_l}=\int_\Theta\left\{\frac{f(x;\theta)\dfrac{\partial^2 f(x;\theta)}{\partial x_h\,\partial x_l}-\dfrac{\partial f(x;\theta)}{\partial x_h}\cdot\dfrac{\partial f(x;\theta)}{\partial x_l}}{f(x;\theta)^2}\right\}\hat{g}(\theta)\,d\theta \tag{24}$$

(by Corollary 5.9 in Bartle [1966] and the regularity conditions on the $f(x;\theta)$). Therefore,

$$\frac{\partial^2 Pl(x,\psi)}{\partial x_h\,\partial x_l}=\frac{1}{2}\frac{\partial^2 L(x)}{\partial x_h\,\partial x_l}\quad\text{at}\quad x=\psi. \tag{25}$$

Now, suppose that there is a local electoral equilibrium at $\psi\in S$. Then, since $\psi\in S$ is a normal point, there exists unique multiplier $\lambda_1,\ldots,\lambda_m$ such that, if we set $L_1(x)=Pl_1(x,\psi)+\lambda_1 g_1(x)+\cdots+\lambda_m g_m(x)$, then $\nabla L_1(\psi)=0$ (e.g., see Theorem 3.2.2 in Hestenes [1975]). Therefore, since $Pl(x,\psi)$ is non-degenerate with respect to the constraints given by (8).

$$L_1''(\psi)=\sum_{l=1}^{n}\sum_{h=1}^{n}\frac{\partial L_1(\psi)}{\partial x_h\,\partial x_l}\cdot w_h\cdot w_l>0$$

for every non-zero vector $w\in R^n$ which satisfies the equation $\nabla g_k(\psi)\cdot w=0$ $(k=1,\ldots,m)$. (For example, see Theorem 3.3.2 in Hesteness [1975].)

Now consider $L(x)$. Since $\nabla L(\psi) = 2 \cdot \nabla Pl(x, \psi)$ at $x = \psi$, we must have $L_2(x) = L(x) + 2 \cdot \lambda_1 g_1(x) + \cdots + 2 \cdot \lambda_m g_m(x)$ implies $\nabla L_2(\psi) = 0$. Additionally, since (25) implies

$$\frac{\partial^2 L(\psi)}{\partial x_h \, \partial x_l} = 2 \cdot \frac{\partial^2 Pl(x, \psi)}{\partial x_h \, \partial x_l} \qquad \text{at} \quad x = \psi,$$

we must have

$$L_2''(\psi) = \sum_{l=1}^{n} \sum_{h=1}^{n} \frac{\partial L_2(\psi)}{\partial x_h \, \partial x_l} \cdot w_h \cdot w_l > 0$$

for every non-zero vector $w \in R^n$ which satisfies $\nabla g_k(\psi) \cdot w = 0$ $(k = 1, \ldots, m)$. Therefore, since $L(x)$ is non-degenerate with respect to the constraints given by (8), ψ must be a strict local maximum of $L(x)$ (e.g., see Theorem 3.3.2 in Hestenes [1975]).

The converse follows similarly. Q.E.D.

Proof of Corollary 3. By the argument of Corollary 2, $L(x)$ is a continuous function of x. Therefore, since S is compact, there is some $\psi \in S$ such that ψ is a local maximum of $L(x)$. By Theorem 3, this ψ must also be a status quo at which there is a local electoral equilibrium. Q.E.D.

ACKNOWLEDGMENTS

We would like to express our thanks for individual comments and suggestions from Ken Arrow, Bob Aumann, Mel Hinich, Bezalel Peleg, Micha Perles, Ariel Rubenstein and an anonymous referee.

REFERENCES

R. Bartle [1965], "The Elements of Integration," Wiley, New York.

L. Cohen and S. Matthews [1980], Constrained plott equilibria, directional equilibria and global cycling sets, *Rev. Econ. Studies* **47**, 975–986.

W. Comaner [1976], The median voter rule and the theory of political choice, *J. Public Econ.* **5**, 169–178.

P. Coughlin and S. Nitzan [1981], Electoral outcomes with probablistic voting and Nash social welfare maxima, *J. Public Econ.* **15**, 113–122.

A. Denzau and A. Kats [1977], Expected plurality voting equilibrium and social choice functions, *Rev. Econ. Studies* **44**, 227–233.

R. Fair [1978], The effect of conomic events on votes for president, *Rev. Econ. Statist.* **60**, 159–173.

P. Fishburn [1975], A probabilistic model of social choice: Comment, *Rev. Econ. Studies* **42**, 197–301.

M. Hestenes [1975], "Optimization Theory," Wiley, New York.

M. HINICH [1977], Equilibrium in spatial voting: The median voter result is an artifact, *J. Econ. Theory* **16**, 208–219.

M. HINICH [1978], The mean versus the median in spatial voting games, *in* "Game Theory and Political Science," (P. Ordeshook, Ed.), New York Univ. Press, New York.

M. HINICH, J. LEDYARD, AND P. ORDESHOOK [1972], Nonvoting and the existence of equilibrium under majority rule, *J. Econ. Theory* **4**, 144–153.

M. HINICH, J. LEDYARD, AND P. ORDESHOOK [1973], A theory of electoral equilibrium: A spatial analysis based on the theory of games, *J. Politics* **35**, 154–193.

M. HIRSCH [1976], "Differential Topology," Springer-Verlag, New York.

M. INTRILIGATOR [1973], A probabilistic model of social choice, *Rev. Econ. Studies* **40**, 553–560.

G. KRAMER [1971], Short-term fluctuation in U.S. voting behavior, 1896–1964, *Amer. Pol. Sc. Rev.* **65**, 131–143.

G. KRAMER [1978], Robustness of the median voter result, *J. Econ. Theory* **19**, 565–567.

G. KRAMER AND A. KLEVORICK [1974], Existence of a "local cooperative equilibrium" in a class of voting games, *Rev. Econ. Studies* **41**, 539–547.

R. LUCE [1959], "Individual Choice Behavior," Wiley, New York.

R. LUCE AND P. SUPPES [1965], Preferences, utility and subjective probability, *in* "Handbook of Mathematical Psychology," Wiley, New York.

S. MATTHEWS [1979], A simple direction model of electoral competition, *Public Choice* **34**, 141–156.

R. MCKELVEY [1975], Policy related voting and electoral equilibria, *Econometrica* **43**, 815–844.

S. NITZAN [1975], Social preference ordering in a probabilistic voting model, *Public Choice* **24**, 93–100.

C. PLOTT [1967], A notion of equilibrium and its possibility under majority rule, *Amer. Econ. Rev.* **57**, 787–806.

P. RAY [1973], Independence from irrelevant alternatives, *Econometrica* **41**, 987–991.

N. SCHOFIELD [1978], Instability of simple dynamic games, *Rev. Econ. Studies* **45**, 575–594.

G. STIGLER [1973], General economic conditions and national elections, *Amer. Econ. Rev.* **63**, 160–167.

JOURNAL OF ECONOMIC THEORY 24, 240–264 (1981)

The Phillips Curve and Bayesian Learning

GEOFFREY LEWIS

Department of Economics, University of Bristol, Alfred Marshall Building,
40 Berkeley Square, Bristol BS8 1HY, England

Received November 13, 1978; revised December 11, 1979

1. INTRODUCTION

The appeal of rational expectations as an equilibrium concept springs from a belief that, with sufficient market experience, rational agents will eventually learn the determinants of equilibrium prices. Despite widespread recognition of its importance, the question of how this learning process takes place has received little attention from economic theorists. For example, in macrotheory, rational expectations are usually regarded as characterising only long-run equilibrium whereas it is problems of short-run adjustment (when agents are engaged in revising their expectations and the allocation decisions which depend on them) that frequently dominate the concern of the policymaker. Yet many macroeconomic models either completely ignore this period of adjustment to equilibrium or, at most, provide a very informal story about it.

The justly celebrated model of Lucas [12] illustrates this tendency to focus exclusively on states of long-run equilibrium. Somehow the agents in his model have acquired a complete knowledge of how their economy works, enabling them to form rational expectations about next period's price and to make associated, efficient allocation decisions. Suppose, to the contrary, one starts by assuming agents do not have such complete knowledge. In particular assume they do not know the values of some non-empty subset of the parameters which characterise the pay-off relevant parts of their environment. The object of this paper is to provide a formal account of how agents learn what these values are by observing market prices and how, contemporaneously, they make allocation decisions that are efficient relative to their current knowledge. If learning proceeds to completion the economy converges to a long-run rational expectations equilibrium.

Pioneering contributions to the theory of learning in simple general equilibrium settings have been made by Cyert and De Groot [3] and particularly Grossman [8, 9]. The approach developed by these writers relies on a subtle synthesis of the theory of rational expectations and Bayesan

0022-0531/81/020240–25$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

learning theory. In this paper we apply a similar approach to introduce learning behaviour into a simple output-inflation model.

Our most interesting achievement is to provide a rigorous choice-theoretic explanation of certain familiar business cycle events in terms of the speed with which private agents learn about and respond to unanticipated changes in the stock of money. The precise way in which the economy's adjustment path depends on agents' current states of knowledge is highlighted. In particular, depending on agents beliefs, there may be a trade-off between output and price stability. However, as agents accumulate observations, they are able to pin down the value of the money stock more and more precisely and in the long-run any Phillip's type trade-off evaporates. That this well-known story is true in our model is not surprising in the light of its structure and assumptions. The interest rather lies in the details of the learning process and in seeing how rational expectations concepts continue to play an important role even out of long-run equilibrium.

In Section 2 we discuss what is the appropriate equilibrium concept to use when agents who are capable of learning about their environment are introduced into an economic model. In Section 3 the basic structure of the model is described. Section 4 develops results which characterise the behaviour of the economy and its agents under various possible states of learning equilibrium and long-run equilibrium. In Section 5 we illustrate how the model generates situations bearing some resemblance to our historical experience and, more recently, lack of experience of output-inflation trade-offs.

2. LEARNING AND EQUILIBRIUM

Prior to presenting our model it is pertinent to ask whether learning behaviour is compatible or not with equilibrium and, if so, with what sort of equilibrium? To discuss the question we distinguish two types of uncertainty: the first consists of any uncertainty which an agent would find impossible or prohibitively expensive to eradicate given his present state of knowledge. Because it arises independently of agents and cannot be modified by any feasible action open to them, it is suitably described as *exogenous* uncertainty. Uncertainty about future weather or demand conditions and the stochastic element in many production processes are examples of exogenous uncertainty. Secondly, there exists the sort of uncertainty which an agent can progressively eliminate as he accumulates observations relevant to the model in terms of which he perceives his environment. As he receives new data the agent behaves as a Bayesian econometrician continually refining his knowledge of unknown parameters. For example, a monopolist facing a stable but unknown demand function learns about its parameters when he sets price and observes sales a number of times. Let us call this a case of *subjective* uncertainty.[1]

Consider two views of the equilibrium status of learning behaviour. Hahn [10] on the one hand, in widening the conventional notion of equilibrium to make it relevant in a world in which imperfect information and learning are present, defines learning in such a way that, essentially, it is a disequilibrium phenomenon. He defines an agent as *learning*, in a sequential economy, when his personal model is not independent of the date t. To be in full informational equilibrium, as Hahn calls it, it is necessary that the policy of each agent, that is, the mapping from the space of all possible messages from his environment to his action space, is independent of t. In effect, full information equilibrium requires the elimination of all subjective uncertainty.

Alternatively, suppose one views learning in a Bayesian framework in which agents treat all unknown parameters as random variables with associated subjective probability distributions. In each period let the agents make allocations which maximise expected utility with respect to all the random variables present be they representing exogenous or subjective uncertainty. Such strategies are efficient only if the distributions of the random variables are *correctly conditioned* on all the information available to agents. But to relate the random variables to his information an agent needs a personal model of how the economy (or the relevant part of it) works. A natural equilibrium condition is to require such a personal model to be "correct" in a rational expectations sense. In particular this personal model should take account of the fact that an important determinant of uncertain future events is the vector of agents' allocation decisions, which in turn depends on their personal models. In other words, efficient learning behaviour implies the self-fulfilling expectations (SFE) condition defined by Grossman [8]. "Economic allocations are equilibrium allocations and economic actors exhibit self-fulfilling expectations if the allocations made by those actors generate price distributions (through market clearing) such that those allocations are optimal with respect to the distributions so generated."

The important point to appreciate here is that this definition of equilibrium is as applicable to learning situations as to Hahn type full-information situations. Since it is often useful to distinguish these, we propose defining two types of equilibrium: long-run full-information equilibrium when only exogenous uncertainty is present and learning equilibrium when both sorts of uncertainty are present. Radner [13] terms the latter momentary equilibria because obviously they are time dependent; agents adapt their allocation decisions to a changing level of information giving rise to a sequence of equilibria. If learning proceeds to the point at which the random variables representing unknown parameters converge in probability to constants, then this sequence converges to a full-information equilibrium. Note that the concept of a learning equilibrium is not trivial; it describes rational, efficient behaviour and it is quite possible to conceive of an agent learning inefficiently through somehow not making the best use of all the information (or all the economic theory!) available to him.

---

[1] In this paper we ignore the possibility that agents can purchase information.

In the model to follow the analysis will proceed in this dual equilibrium framework. We shall study the comparative dynamics of the time paths of output and prices as the sequence of momentary, learning equilibria converges towards the long-run, full-information equilibrium.

3. THE STRUCTURE OF THE MODEL

Our model is based on the previously mentioned model of Lucas [12]. We now give a brief description of its basic elements.

At any time the population of agents is $2N$. Each agent lives for 2 periods, N agents being born and N dying at the beginning of each period. For instance, in period t, there are N younger generation or age-0 agents who were born at the beginning of period t and N older generation or age-1 agents, who were born at the beginning of period $t-1$. Within a generation, all agents are identical in every respect so that we can formulate the model in terms of a representative agent from each generation.

In each period transactions occur between the older and the younger generations. Only the younger generation is productive. It produces a single good by means of labour inputs such that r units of labour produce r units of output, for all non-negative r. The good is perishable so the only way an age-1 agent can consume is by purchasing output from the younger generation. This he does by offering money he has acquired through selling a proportion of his output in the previous period. The total quantity of money in the economy, M, is generally fixed but may be changed from time to time by the monetary authority. In a learning phase, age-0 agents are uncertain about the true value of M or, equivalently, the average amount of money per member of the older generation, $m = M/N$. In a Bayesian fashion, they regard m as a random variable and their beliefs about its value are held in the form of a prior probability distribution. Assume m belongs to a compact subset, Ω, of the positive real numbers. When the true value of m is known, it is, of course, not a random variable and we shall denote it m^0.

Lack of information about m constitutes the source of subjective uncertainty in the economy. The source of exogenous uncertainty lies in the existence of two distinct markets between which there is no communication. In each period, age-0 agents are allocated between these two markets according to a random variable θ; the proportion $\theta/2$ of them to one and

$1-(\theta/2)$ to the other. θ is assumed to be continous and symmetrically distributed about one on Θ, a closed subset of $(0, 2)$. Denote its probability density function (p.d.f.) by $g(\cdot)$.

Exactly half the total amount of money circulates in each market. Trading in each market is perfectly coordinated by means of a Walrasian auctioneer, or some such device, so that all transactions take place at a single market-clearing price which of course will generally differ between the markets.

The preferences of agents are such that each age-1 agent supplies his total cash holdings inelastically in exchange for the consumption good. Each age-0 agent has preferences for current consumption, current leisure and future consumption according to the utility function

$$U(c_t, n_t) + E\{V(c_{t+1})\}, \tag{1}$$

where

c_t is consumption in period t,
n_t is labour supplied (output) in period t,
E is the expectation operator.

U and V are assumed to be strictly concave, continuously twice differentiable and to have normally signed first derivatives. Also

$$\lim_{c\to 0} V'(c) = \infty; \qquad \lim_{c\to 0} V'(c) = 0 \tag{2}$$

and

$$V''(c)c + V'(c) > 0. \tag{3}$$

To form efficient expectations about the future price of the consumption good, p_{t+1}, an age-0 agent must condition his subjective probability distribution concerning p_{t+1} on the observable current price, p_t, because the latter is a signal related to the unknown parameter m. Denote this distribution function initially by $F(p_{t+1} \mid p_t)$. Later we shall characterise it more explicitly showing how the agent forms the price expectations represented by this distribution. Denoting the money received by an age-0 agent in exchange for his output by y_t, his decision problem is

$$\max_{c_t, n_t, y_t} U(c_t, n_t) + \int V(y_t/p_{t+1})\, dF(p_{t+1} \mid p_t) \tag{4}$$

subject to the budget constraint

$$p_t(n_t - c_t) - y_t \geqslant 0.$$

In a manner very similar to Lucas, it can be shown that the solutions for c_t, n_t and $p_t\pi$, where π is the Lagrange multiplier, can be expressed as functions of y_t/p_t the quantity of real balances purchased: n_t and $p_t\pi$ are increasing functions of y_t/p_t while c_t is a decreasing function. The first-order condition which equates the marginal cost of acquiring cash (in terms of foregone period t utility) to its marginal benefit in the form of greater expected utility in period $t + 1$ is, assuming $y_t > 0$:

$$\int V'\left(\frac{y_t}{p_{t+1}}\right)\frac{1}{p_{t+1}}\,dF(p_{t+1}\mid p_t)-\pi=0. \tag{5}$$

Denoting the marginal cost of holding real balances, $p_t\pi$, as a function of y_t/p_t by $h(y_t/p_t)$, (5) can be written

$$h\left(\frac{y_t}{p_t}\right)\frac{1}{p_t}=\int V'\left(\frac{y_t}{p_{t+1}}\right)\frac{1}{p_{t+1}}\,dF(p_{t+1}\mid p_t). \tag{6}$$

Implicitly (6) defines a demand for money function in terms of p_t and the distribution of future prices.

4. EXPECTATIONS AND LEARNING

In this section we first derive a restriction on the period t price, given age-0 agents future price expectations conditional on p_t, which is necessary to ensure that the market clears in period t. Then we address the important question: How does an agent form his price expectations and revise them as he receives further information? We define momentary and long-run equilibria and we show how the variables of our learning model can be interpreted in such a way that it is formally equivalent to Lucas' full-information equilibrium model.

4.1. *Money—Supply and Demand*

Since the two markets in the model are not linked in any way, certain aspects of the learning adjustment path and the long-run equilibrium solution can be determined separately in each. We shall confine our attention initially to one of the markets—that which receives $\theta/2$ of the younger generation. Then it is a simple exercise to substitute $2 - \theta$ for θ to obtain the equilibrium solution for the other market.

Recall that age-1 agents supply all their cash holdings inelastically. To ensure that the total money demanded in period t equals the total supplied therefore requires that

$$\frac{\theta_t N y_t}{2} = \frac{Nm}{2} \quad \text{or} \quad y_t = \frac{m}{\theta_t}. \tag{7}$$

Since $m/\theta_t > 0$, y_t must be positive in equilibrium so, substituting for y_t, the marginal condition (6) becomes

$$h\left(\frac{m}{\theta_t p_t}\right)\frac{1}{p_t} = \int V'\left(\frac{m}{\theta_t p_{t+1}}\right)\frac{1}{p_{t+1}}\, dF(p_{t+1} \mid p_t). \tag{8}$$

A value of p_t and an expectational distribution over p_{t+1} conditional on p_t which satisfy this equation would induce age-0 agents to demand just that quantity of money that is supplied by the age-1 agents: the market would clear.

4.2. *The Price Function*

Each age-0 agent is assumed to form his price expectations by means of a personal model of the economy. This consists essentially of a function which maps states of the economy into market-clearing prices. A *state of the economy* at time t, s_t, is a vector specifying a value of m, a p.d.f. $k_t(\cdot)$, describing age-0 agents' beliefs about m at t and a realization of θ at t. Let K be the set of all possible subjective p.d.f.'s on Ω. Then $s_t = (m, k_t, \theta_t)$, $s_t \in S$, where $S = \Omega \times K \times \Theta$.

We define a *price function* as a mapping,

$$p\colon S \to R^+,$$

which describes agents' understanding of a regular relationship between the underlying state of the economy and the market price.

On observing a price in the current period, p_t, an age-0 agent ascertains the inverse image of p_t under the price function and infers as much as possible about m. This results in his forming a posterior p.d.f. on m, $k_{t+1}(\cdot)$ and, again via the price function $p(m, k_{t+1}, \theta_{t+1})$, price expectations on p_{t+1}. Facing p_t and these expectations conditioned on p_t, each age-0 agent formulates his demand for cash balances, y_t. If (i) the market clears and (ii) expectations are "fulfilled" in the sense that $p(m, k, \theta)$ *correctly* describes the determination of market-clearing price then the consequence is a *learning* or *momentary equilibrium.*

4.3. *Momentary and Long-run Equilibria*

Formally, *a momentary equilibrium is a non-degenerate subjective p.d.f. on Ω, $k_t(\cdot)$ and a price function $p^*(m, k_t, \theta_t)$ such that*

$$h\left(\frac{m}{\theta_t p^*(m, k_t, \theta_t)}\right)\frac{1}{p^*(m, k_t, \theta_t)}$$

$$= \int V'\left(\frac{m}{\theta_t p^*(m, k_{t+1}, \theta_{t+1})}\right)\frac{1}{p^*(m, k_{t+1}, \theta_{t+1})} g(\theta_{t+1})\, k_{t+1}(m)\, d\theta_{t+1}\, dm \tag{9}$$

for all $(m, k_t, \theta_t) \in S$, *where* k_{t+1} *is the posterior p.d.f. of* m *conditioned by the current observation* $p_t = p^*(m, k_t, \theta_t)$.

If prices have been observed a sufficient number of times and the learning process converges, the true value of m will become fully known to agents. The learning process converges if the sequence of p.d.f.'s $(k_0, k_1, k_2 \cdots k_t \cdots)$ converges to a degenerate p.d.f. with all its probability mass concentrated at m^0, the true value of m. When this happens, the economy attains long-run equilibrium. Formally, *a long-run or full information equilibrium is a value of* m, m^0, *and a price function* $p^*(\theta_t; m^0)$: $\Theta \to R^+$ *such that*

$$h\left(\frac{m}{\theta_t p^*(\theta_t; m^0)}\right)\frac{1}{p^*(\theta_t; m_0)}$$

$$= \int V'\left(\frac{m}{\theta_t p^*(\theta_{t+1}; m^0)}\right)\frac{1}{p^*(\theta_{t+1}; m^0)} g(\theta_{t+1})\, d\theta_{t+1} \quad \textit{for all } \theta_t \in \Theta \tag{10}$$

Equation (10) is obtained from (6) by substituting $p^*(\theta; m^0)$ for p_t and p_{t+1}. It should be noted that both these types of equilibrium are of the SFE variety. Much of the remainder of the paper is devoted to exploring the properties of these equilibrium price functions.

4.4. *The Evolution of Expectations*

If $p^*(m, k_t, \theta_t)$ were one-to-one in m, given k_t, a single observation of the market clearing price would reveal the unknown m and the economy would move in one step to a full information equilibrium.[2] But the central feature of this economy is that money and real signals are mixed in such a way that $p^*(m, k_t, \theta_t)$ is not one-to-one in m. In fact p^* reveals the *ratio* m/θ_t for given k_t, but no more. To understand this, one need only appreciate that for m to be revealed by p^*, it would be necessary that some agents (i) know the true m and (ii) change their behaviour (supply or demand) when m changes, independently of their response to changes in θ_t. But neither type of agent knows about m beyond what is already incorporated in k_t. The equilibrium cash transaction per age-0 agent is m/θ_t and this, *at most*, is what p^* can reveal to an age-0 agent. In fact p^*, if it exists, does reveal m/θ_t since it is easy to prove from (9) and from the fact that $h(\cdot)$ is strictly increasing and $V'(\cdot)$ is strictly decreasing that $p^*(m, k_t, \theta_t)$ is one-to-one in m/θ_t.[3]

---

[2] By one-to-one in m given k_t we mean $m_1 \neq m_2 \Rightarrow p^*(m_1, k_t, \theta_t^1) \neq p^*(m_2, k_t, \theta_t^2)$ for all $\theta_t^1, \theta_t^2 \in \Theta$.

It follows from this one-to-one property that the posterior p.d.f. of m, $k_{k+1}(\cdot)$, can be conditioned equivalently on p^* or on the corresponding m/θ_t. It is obtained by Bayes' rule in the normal way, thus

$$k_{t+1}(m) = \frac{\xi(m/\theta_t \mid m)\, k_t(m)}{\int_\Omega \xi(m/\theta_t \mid m)\, k_t(m)\, dm}, \tag{11}$$

where $\xi(\cdot \mid m)$, the conditional p.d.f. of the random variable m/θ_t given m, is derived from $g(\cdot)$.

With the decidedly miserable allotted life-span of only two periods an agent clearly has rather limited learning opportunities. Although one could extend the model to admit multiperiod life-cycles it would be at the cost of considerable complexity which it seems desirable to avoid for the moment. So to allow a reasonable number of observations on which learning can be based, we make the not entirely unrealistic assumption that agents pass on their knowledge of m to their descendants—they bequeath their priors. As successive generations observe prices, intuitively one would expect the posterior p.d.f. to become more and more concentrated about the true value of m. Such convergence can be guaranteed provided certain regularity conditions concerning the family of p.d.f.'s $\{\xi(\cdot \mid m),\ m \in \Omega\}$ are satisfied.[4] For example if $k_t(\cdot)$ is normal (and for large enough t, all posterior p.d.f.'s associated with regular families $\{\xi(\cdot \mid m),\ m \in \Omega\}$ are approximately normal)[5] then the expected value of m under k_t (denoted m_t) tends to the true value m^0 as observations accumulate, i.e., with probability one,

$$\lim_{t \to \infty} m_t = m^0.$$

4.5. *Lucas Equivalence*

In order to establish an important formal equivalence between our learning model and Lucas' full information model, it is helpful to express m in terms of its current expected value, m_t. Define

$$x_t = m/m_t, \tag{12}$$

where $x_t \in X$, derived from Ω.

One can think of x_t as a measure of the extent that age-0 agents under or over-estimate m. Also it is sometimes helpful and valid to think of m_t as the maximum likelihood estimator of m^0. The p.d.f. of x_t, denoted $f_t(\cdot)$, is derived from $k_t(\cdot)$ and clearly $E[x_t] = 1$ for all t. Let F be the set of all possible p.d.f.'s, on X.

---

[3] Assume $p^*(m', k_t, \theta'_t) = p^*(m'', k_t, \theta''_t)$, where $m'/\theta'_t \neq m''/\theta''_t$ and obtain a contradiction.

[4] See DeGroot [5, Chap. 10].

Describing the current state of knowledge about m by the pair $\{m_t, f_t(\cdot)\}$ instead of by $k_t(\cdot)$ enables one to consider the effect of a *fully anticipated change* in m on the equilibrium price p^*. For instance, one can imagine an agent becoming aware that m has increased by 10% without his subjective beliefs becoming any more "concentrated." His state of knowledge changes from $\{m_t, f_t(\cdot)\}$ to $\{1.1\ m_t, f_t(\cdot)\}$. Because agents completely lack money illusion it follows that such a fully anticipated increase will simply raise equilibrium prices by 10% but have no real effects.

These remarks suggest a specification of the price function which incorporates both the monotonicity (one-to-one) property with respect to m/θ_t and the neutrality property of fully anticipated increases in m. Consider,

$$p_t = m_t \phi\left(f_t, \frac{m}{m_t \theta_t}\right). \tag{13}$$

(Check: p_t is homogeneous of degree zero in m and θ_t and homogeneous of degree one in m and m_t.) Define $z_t = x_t/\theta_t$, where $z_t \in Z$. Formally, $\phi(f_t, z_t)$ is a function ϕ: $F \times Z \rightarrow R^+$ and $\phi^*(f_t, z_t)$, the momentary equilibrium value of $\phi(f_t, z_t)$, is defined analogously to $p^*(m, k_t, \theta_t)$. The long-run equilibrium function $\phi^*(\theta_t)$ is defined analogously to $p^*(\theta_t; m^0)$.

Substituting $m_t\phi^*$ for p^* in (9) and using $m_t x_t = m_{t+1} x_{t+1} = m$ one obtains

$$h\left(\frac{z_t}{\phi^*(f_t, z_t)}\right)\frac{z_t}{\phi^*(f_t, z_t)} = \int V'\left(\frac{\phi_{t+1}}{\phi_t}\frac{z_{t+1}}{\phi^*(f_{t+1}, z_{t+1})}\right)$$
$$\frac{\theta_{t+1} z_{t+1}}{\theta_t \phi^*(f_{t+1}, z_t)} \hat{H}_t(\theta_t \mid z_t)\, H_{t+1}(\theta_{t+1}, z_{t+1})\, d\theta_t\, d\theta_{t+1}\, dz_{t+1}, \tag{14}$$

where $\hat{H}_t(\cdot \mid z_t)$ is the p.d.f. of θ_t conditional on z_t and $H_t(\cdot, \cdot)$ is the joint p.d.f. of z_t and θ_t. $\hat{H}_t$ and H_t are both derived from $g(\cdot)$ and $f_t(\cdot)$.

Equation (14) is formally identical to the central equilibrium relationship in Lucas' model.[6] Although his model includes no parameter uncertainty and no learning it contains an additional source of exogenous uncertainty which affects the money supply. It is this that gives rise to the formal similarity: at a given time there is no difference between subjective and exogenous uncertainty. But over time an agent can reduce his subjective uncertainty by learning. This difference shows up in (13). The function ϕ has the current subjective p.d.f. of x_t, $f_t(\cdot)$, as an argument. In the equivalent Lucas function, the corresponding p.d.f. stays constants over time so it would be gratuitous to include it as an argument.

4.6. *Alternative Representations of Subjective Probability*

If the subjective p.d.f.'s $k_t(\cdot)$, $t = 0, 1, 2,\ldots$ belong to what is known in Bayesian analysis as a *conjugate family*,[7] it may be simpler and more useful to represent the state of knowledge about m by a vector of parameters which picks out a member of the family rather than by the p.d.f. itself. For example, if the conjugate family were a family of normal p.d.f.'s then the two element vector (m_t, τ_t), where τ_t is the precision[8] of m, is sufficient to completely describe the period t state of knowledge about m. In such a case one can rewrite $\phi(f_t, z_t)$ as a function of μ_t and z_t, where μ_t $(=m_t^2\tau_t)$ is the precision of x_t. Define

$$\psi(\mu_t, z_t) \equiv \phi(f_t(\cdot\,; \mu_t), z_t). \tag{15}$$

More generally ψ would have as arguments as many parameters as necessary to specify the p.d.f. $f_t(\cdot)$. Using this sort of specification it may be possible to derive interesting comparative static effects of changes in the parameters on the sequence of momentary equilibria. We derive such a result below when, as in (15), a member of the family of random variables x_t, $t = 0, 1, 2,\ldots$ can be specified solely by its precision.

5. SOME THEOREMS

Let us first restate the Lucas theorems appropriately modified to fit the learning context of the present model.[9]

THEOREM 1. *Equation* (19) *has exactly one continuous solution* $\phi^*(f, z)$ *on* $F \times Z$ *with* $z/\phi^*(f, z)$ *bounded. The function* $\phi^*(f, z)$ *is strictly positive and continuously differentiable. Further,* $m_t\phi^*(f_t, z_t)$ *is the unique momentary equilibrium price function.*

THEOREM 2. *Suppose* $\theta = 1$ *with probability one. Let* w^* *be the unique solution to*

$$h(w) = V'(w).$$

---

[7] When for whatever outcome the posterior p.d.f. of an unknown parameter belongs to the same family as the prior p.d.f. the family is called a *conjugate family*. See DeGroot [5, Chap. 9].

[8] The precision of a random variable is the reciprocal of its variance.

[9] The proofs are straightforward modifications of Lucas' proofs. The major adaptation required is to the proof of Theorem 1, where the contraction mapping is applied to the space of all bounded, continuous functions on $F \times Z$ rather than on $(-\infty, \infty)$. A suitable metric on F must be chosen.

Then $p^*(m_0) = m_0\phi^*(1) = m_0/w^*$ *is the unique long-run equilibrium price "function." No momentary equilibrium exists.*

THEOREM 3. *Suppose* $x = 1$ *with probability* 1. *Then* $p^*(\theta_t; m^0) = m^0\phi^*(\theta_t)$ *is the unique long-run equilibrium price function.* $\phi^*(\theta)$ *is a continuously differentiable function with an elasticity between zero and one.*

THEOREM 4. *Suppose the random variables* x_t *and* θ_t *are distributed such that, first, for any fixed* $\bar{\theta} \in \Theta$, $pr\{\theta \leqslant \bar{\theta} \mid x_t/\theta_t = z_t\}$ *is an increasing function of* z_t, *and, second, that for any fixed* $\bar{x} \in X$, $pr\{x_t \leqslant \bar{x} \mid x_t/\theta_t = z_t\}$ *is a decreasing function of* z_t. *Then* $p^*(m, k_t, \theta_t) = m_t\phi^*(f_t, z_t)$ *is the unique momentary equilibrium price function.* ϕ^* *is a continuously differentiable function with an elasticity with respect to its second argument between zero and one.*

Theorem 1 establishes the existence and uniqueness of the momentary equilibrium price function. Theorem 2 applies when there is no exogenous uncertainty in the model. Agents lack information only concerning the value of m. But this is completely revealed by the current market clearing price. Even if m_0 is changed agents are aware that all price variations are purely monetary and since they do not suffer money illusion their demand and supply behaviour is completely inelastic with respect to such changes. In consequence, the equilibrium price is exactly proportional to the quantity of money. Learning is unnecessary and forming correct expectations of next period's price is a trivial exercise.

Theorem 3 is the polar opposite situation to that of Theorem 2. There is no subjective uncertainty, agents have perfect information about m. ($m_t = m^0$ for all t.) But variations in θ are a source of exogenous uncertainty which has real effects. A higher market-clearing price in the current period unambiguously indicates to younger generation agents a better relative price for their current output (equivalently, a lower price for their future consumption). The agents respond by increasing output and decreasing current consumption to acquire higher real balances. This increased demand for future consumption means that, to clear the market, the current price must rise less than proportionally to any increase in $1/\theta$.

Theorem 4 applies to the situation in which we are particularly interested, when both subjective and exogenous uncertainty are present. The restrictions on the distributions of x_t and θ_t will be obeyed by a wide class of p.d.f.'s. They are made to rule out exceptions which can be constructed rather easily, but are of little practical relevance. The most interesting and important implication of the theorem is that subjective uncertainty about the money supply has real effects. For example, suppose x_t were to increase owing to the monetary authority secretly increasing m or simply due to an "unfortunate" previous price observation leading to a revision of m_t away from, rather than towards, m. Then the equilibrium quantity of real balances demanded by each age-0 agent $y_t/p_t = z_t/\phi^*(f_t, z_t)$ will increase with an associated increase and decrease in n_t and c_t, respectively.

We need a further theorem to establish the effects of a ceteris paribus change in the subjective beliefs of age-0 agents. This is not considered by Lucas since the p.d.f. corresponding to $f_t(\cdot)$ remains constant in his model. But here f_t generally becomes more concentrated as learning proceeds and one can also envisage various exogenous events which result in a greater or smaller degree of confidence among younger generation agents concerning their beliefs and expectations. We consider the special case where the p.d.f.'s $(f_0, f_1, f_2, \ldots, f_t, \ldots)$ comprise a conjugate family in which each member is fully specified by a single parameter, namely, the precision of x_t, μ_t. First we give an intuitive explanation of the effect of a small change in μ_t on momentary equilibrium: as the precision with which an age-0 agent believes the true m lies in relation to m_t increases, the more he will interpret a, say, high observed current price as being the result of a low θ_t rather than a high x_t. Since this means a shift in the intertemporal terms of trade in favour of the future his demand for real balances will increase. But the actual quantity of money available for distribution per age-0 agent remains the same. So to choke off the increase in production which age-0 agents will now wish to supply in exchange for money, requires a *lower* current price. The comparative static result then is that the momentary equilibrium price is a decreasing function of μ_t for "high" values of z_t (or p_t) and an increasing function for "low" z_t values. The latter half of the result follows because low z_t values will be interpreted as more likely to have arisen from a high θ_t than a low x_t, the greater μ_t (i.e., the smaller the variance of x_t). The dividing point between high and low values of z is its expected value, namely, one.

Unfortunately matters are not as straightforward as this account suggests. An additional influence is at work because when p.d.f.'s are revised in a Bayesian manner, an exogenous change in μ_t causes a similar change in μ_{t+1} (e.g., in the case of a conjugate family of normal p.d.f.'s $d\mu_{t+1}/d\mu_t = 1$). Now a higher μ_{t+1} has little direct effect on the *mean value* of anticipated consumption in period $t+1$ but it does decrease its variability. How much this affects the marginal expected utility of money saved for future consumption depends on agents' attitude to risk.

We now formalise this discussion. Distinguish the two effects of $d\mu_t$ on the marginal expected utility of money discussed above by calling the first the *mean effect* and the second the *variance effect*. To indicate precisely what we mean by these terms substitute $m_t\psi^*(\mu_t, z_t)$ for p_t^* in (9). Since a ceteris paribus change in μ_t leaves m/θ_t unaltered, let m/θ_t be denoted simply as y, a constant. We obtain,

$$h\left(\frac{y}{m_t\psi^*(\mu_t, z_t)}\right)\frac{y}{m_t\psi^*(\mu_t, z_t)}$$
$$= \int V'\left(\frac{y}{m_{t+1}(\mu_t)\,\psi^*(\mu_{t+1}, z_{t+1})}\right)\frac{y}{m_{t+1}(\mu_t)\,\psi^*(\mu_{t+1}, z_{t+1})}$$
$$\tilde{H}_{t+1}(z_{t+1};\mu_{t+1})\,dz_{t+1}, \tag{16}$$

where $\tilde{H}_t(\cdot;\mu_t)$ is the marginal p.d.f. of z_t. Denoting $y/[m_{t+1}(\mu_t)\,\psi^*(\mu_{t+1}, z_{t+1})]$ by $c(\mu_t, z_{t+1})$, differentiating the R.H.S. of (16) w.r.t. μ_t gives

$$\frac{-y}{m_{t+1}}\int [V''(c)c + V'(c)]\left[\frac{m'_{t+1}(\mu_t)}{m_{t+1}\psi^*} + \frac{\psi^*_{\mu_{t+1}}}{\psi^{*2}}\frac{d\mu_{t+1}}{d\mu_t}\right]$$
$$\tilde{H}_{t+1}(z_{t+1};\mu_{t+1})\,dz_{t+1} + \int V'(c)c\,\frac{\partial\tilde{H}_{t+1}}{\partial\mu_{t+1}}\frac{d\mu_{t+1}}{d\mu_t}\,dz_{t+1}. \tag{17}$$

The *mean effect* of a small change in μ_t is defined as

$$-y\,\frac{m'_{t+1}(\mu_t)}{m^2_{t+1}}\int [V''(c)c + V'(c)]\,\frac{1}{\psi^*}\,\tilde{H}_{t+1}(z_{t+1};\mu_{t+1})\,dz_{t+1}, \tag{18}$$

while the *variance effect* is the difference between (17) and (18).

We can now make some precise statements about the effects of changes in μ_t on momentary equilibrium. Theorem 5 proves that if the mean effect "dominates" the variance effect and if a certain reasonable restriction on the random variables x_t (or m) and θ_t holds, then ψ^*_μ behaves as we initially suggested in our intuitive story. A discussion following the theorem aims to show that the dominance of the mean effect is likely, but unfortunately we are unable to offer a proof. We also conjecture a result about ψ^*_μ when z_t is "close" to one.

THEOREM 5. *Suppose the random variables x_t and θ_t are distributed such that for any fixed $\bar{x} \in X$ and $\bar{z} \in Z$, $P\{x \leqslant \bar{x} \mid x_t/\theta_t = \bar{z}\}$ is an increasing function of μ_t if $\bar{z} > 1$ and a decreasing function if $\bar{z} < 1$. Then, for given z_t, if the absolute value of the mean effect is greater than the absolute value of the variance effect, $\psi^*(\mu_t, z_t)$ decreases with respect to μ_t if the given $z_t > 1$ and increases if $z_t < 1$.*

Proof. The result follows from (17) and the fact that $m'_{t+1}(\mu_t)$ is negative if $z_t > 1$ and positive if $z_t < 1$ which we now demonstrate. Since $x_t = m/m_t$, $P\{m \leqslant \bar{m} \mid x_t/\theta_t = \bar{z}\}$ is the same as $P\{x \leqslant \bar{x} \mid x_t/\theta_t = \bar{z}\}$ where $\bar{m} = m_t\bar{x}$. The conditional p.d.f. of m given z_t is $k_{t+1}(\cdot)$.

$$\therefore\quad P\{m \leqslant \bar{m} \mid z_t\} = \int_0^{\bar{m}} k_{t+1}(m)\,dm = K_{t+1}(\bar{m}; z_t, \mu_t),$$

where K_{t+1} is the distribution function corresponding to k_{t+1} (its dependence on z_t and μ_t are shown explicitly). $\therefore$ the restrictions on x_t and θ_t are equivalent to

$$\frac{\partial K_{t+1}}{\partial \mu_t} \begin{cases} > 0 & \text{if} \quad z_t > 1 \\ < 0 & \text{if} \quad z_t < 1. \end{cases} \tag{19}$$

The mean of m in period $t+1$ is

$$m_{t+1}(\mu_t) = \int_\Omega m k_{t+1}(m)\, dm.$$

Integrating the R.H.S. by parts,

$$m_{t+1}(\mu_t) = [mK_{t+1}]_\Omega - \int_\Omega K_{t+1}(m; z_t, \mu_t)\, dm,$$

noting that the first right hand term is a constant. Differentiating w.r.t. μ_t and using (19) we have

$$\begin{aligned} m'_{t+1}(\mu_t) &< 0 \qquad \text{if} \quad z_t > 1 \\ &> 0 \qquad \text{if} \quad z_t < 1. \end{aligned}$$

We can now sign the mean effect. Since by (3) $V''(c)c + V'(c) > 0$, the mean effect is positive if $z_t > 1$ and negative if $z_t < 1$. Differentiating the L.H.S. of (16) w.r.t. μ_t gives

$$-\psi^*_{\mu_t}\left(h' \frac{y}{m_t \psi^*} + h\right)\frac{y}{m_t \psi^*}.$$

Since h and h' are positive functions and the mean effect is assumed to dominate the variance effect in absolute value, the required result follows immediately. Q.E.D.

If $\psi^*_{\mu_{t+1}}$ behaves according to the theorem it seems likely that the mean effect will dominate the variance effect. For the mean effect is the probability sum (integral) of the terms $\{V''[c(\mu_t, z_{t+1})]\, c(\mu_t, z_{t+1}) + V'[c(\mu_t, z_{t+1})]\}\, \psi^{*-1}$ over $z_{t+1} \in Z$ multiplied by $-y_t m'_{t+1}(\mu_t)/m^2_{t+1}$. By contrast, part of the variance effects consists of a weighted sum of the *same* terms but each weight, $-(y_t \psi^*_{\mu_{t+1}}/m_{t+1}\psi^*)(d\mu_{t+1}/d\mu_t)\, \bar{H}_{t+1}\, dz_{t+1}$ is positive if $z_{t+1} > 1$ and negative if $z_{t+1} < 1$. The cancellation of oppositely signed terms will make the absolute value of this part of the variance effect smaller than the absolute value of the mean effect, if $m'_{t+1}(\mu_t)$ is large enough. By continuity since $m'_{t+1}(\mu_t) = 0$ when $z_t = 1$ this will be when z_t is not "too close" to one.

Consider the second part of the variance effect:

$$\int V'(c)c\frac{\partial \bar{H}_{t+1}}{\partial \mu_{t+1}}(z_{t+1};\mu_{t+1})\frac{d\mu_{t+1}}{d\mu_t}\,dz_{t+1}.$$

This is the rate of change of the marginal expected utility of money when the random variable x_t (and hence z_t) becomes more concentrated.[10] If z_{t+1} is distributed so that a decrease in its precision is equivalent to a mean-preserving spread in the sense defined by Rothschild and Stiglitz [14] then the sign of this term is negative (positive) if $V'(c)c$ is a convex (concave) function of z_{t+1}.[11] This in turn depends on the sign of a function of the second and third derivatives of $V(c)$ and the first and second partial derivatives of ψ^* w.r.t. z_{t+1}. One would like a clear result relating the convexity (concavity) of $V'(c)c$ to the degree of risk aversion in the second period and whether it is increasing or decreasing with consumption. However, at the level of generality with which we have specified $V(c)$ and the p.d.f.'s of x_t and θ_t, the sign of the second part of the variance effect is indeterminate indicating that, at least for a range of cases, it will also be sufficiently small to be dominated by the income effect.

What makes it difficult to obtain precise conditions under which the sum of the two effects is unambiguously signed is the SFE equilibrium nature of the model: a change in μ_t affects not only the distribution of z_{t+1} but also, for given z_{t+1}, the utility of a unit of money, through the equilibrium price. In contrast, most studies of behaviour when the degree of uncertainty is changing have used (understandably) a more tractable partial equilibrium approach.

When z_t is close to one, $m'(\mu_t)$ is small so the variance effect may dominate. Since, as we have seen, its sign depends on the third derivative of $V(c)$ which is left unspecified by our assumptions, the demand for money

---

[10] A more concentrated r.v. could mean either one that has greater precision or one that is less risky. As is now well known these are not always the same; see Rothschild and Stiglitz [14]. However, for families of r.v.'s whose p.d.f.'s differ only by a location shift, such as normal p.d.f.'s, they are the same.

[11] *Proof.* Integrate $\int V'(c)c(\partial \bar{H}_{t+1}(z_{t+1};\mu_{t+1})/\partial\mu_{t+1})\,dz_{t+1}$ by parts, twice, omitting subscripts and letting $\Gamma(\cdot)$ be the distribution function of z_{t+1}:

$$\int V'(c)c\frac{\partial \bar{H}}{\partial \mu}\,dz = -\int \frac{\partial}{\partial z}[V'(c)c\frac{\partial \Gamma}{\partial \mu}\,dz$$

$$= \int \frac{\partial^2[V'(c)c]}{\partial z^2}\left\{\int_0^z \frac{\partial \Gamma}{\partial \mu}(s;\mu)\,ds\right\}dz.$$

Diamond and Stiglitz [6] have shown that if decreasing μ corresponds to a mean-preserving spread in the distribution of z then the term in curly brackets is negative. If $V'(c)c$ is convex (concave) in z, the other term is everywhere positive.

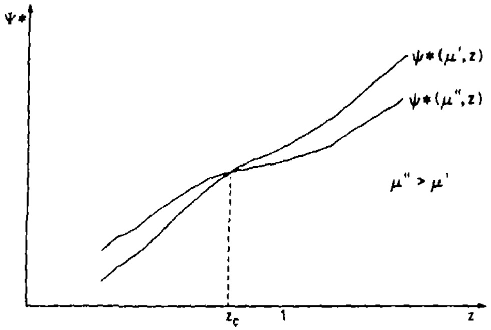


FIG. 1. Effect of changing confidence on the equilibrium price function.

may either increase or decrease when μ_{t+1} increases, holding prices and expectations constant. So ψ^*_μ correspondingly must be negative or positive to restore equilibrium. But, as z_t departs from unity, at some critical point $z_c < (>)1$, we conjecture the mean effect will become dominant and $\psi^*_\mu > (<)0$ for all $z \leqslant (\geqslant z_c)$. The former case is sketched in Fig. 1.

6. COMPARATIVE LEARNING DYNAMICS

We can now use the model to explore the interactions between monetary expansion or contraction, prices and real output. The three are connected through the actions of agents who are rational, imperfectly informed but capable of learning in the manner of Bayesian statisticians. We have already seen that when both subjective and exogenous sources of uncertainty are present, observing the current price does not enable an agent to distinguish their separate influence. Specifically consider a $100r\%$ increase in $x_t = m/m_t$, the measure of subjective uncertainty concerning a purely *monetary* matter. This will result in exactly the same change in the observed momentary equilibrium price as a $(100r/(1+r))\%$ decrease in θ_t, the measure of exogenous uncertainty which concerns a *real* event.

Such changes will also have equal effects on the current real output and consumption behaviour of age-0 agents. Recall that current supply and consumption can be expressed as increasing and decreasing functions respectively of the amount of real balances demanded, y_t/p_t^*. But in momentary equilibrium $y_t/p_t^* = m/\theta_t p_t^* = x_t/\{\theta_t \psi^*(\mu_t, x_t/\theta_t)\}$. So we can define a momentary equilibrium employment function $n(\mu_t, x_t/\theta_t)$ or $n(\mu_t, z_t)$ where, if the assumptions made in Theorems 4 and 5 hold,

$$\frac{\partial n}{\partial z_t} > 0; \qquad \frac{\partial n}{\partial \mu_t} \begin{cases} >0 & \text{if } z_t > 1 \\ <0 & \text{if } z_t < 1. \end{cases} \tag{20}$$

It is clear now how the model generates a Phillips curve type trade-off between real output and demand expansion or contraction. The results above apply to the market which receives fraction $\theta_t/2$ of the younger generation. If we assume that the $1-\theta_t/2$ of age-0 agents who enter the other market have the same beliefs about m at time t then their supply behaviour and the determination of the equilibrium price is given by $n(\mu_t, x_t/2-\theta_t)$ and $\psi^*(\mu_t, x_t/2-\theta_t)$, respectively. The total output of the economy, Y_t, is

$$Y_t = \frac{1}{2}\theta_t Nn\left(\mu_t, \frac{x_t}{\theta_t}\right) + \frac{1}{2}(2-\theta_t)Nn\left(\mu_t, \frac{x_t}{2-\theta_t}\right). \tag{21}$$

If agents have perfect information about m (i.e., $x_t = 1$, $\mu_t = \infty$) then per capita output varies in each market with θ_t. But because higher per capita output occurs in the market with fewer producers, total output will tend to be fairly constant from period to period. Y_t, in this situation, can be identified as the "natural rate" of output, Y_t^n. If $n(\theta_t)$ is the long run equilibrium supply function corresponding to $\psi^*(\theta_t)$ then

$$Y_t^n = \tfrac{1}{2}\theta_t Nn(\theta_t) + \tfrac{1}{2}(2-\theta_t)Nn(2-\theta_t). \tag{22}$$

Clearly Y_t^n is invariant with respect to whatever changes in m occur as long as they are *fully and correctly anticipated*. And when they are, of course, the price level adjusts in full proportion to the change. This inelasticity of output with respect to fully anticipated changes in monetary conditions can be interpreted as the much discussed verticality of the long-run Phillips curve.

When monetary changes are not perfectly anticipated, as measured by x_t, then from (20) it is clear there will be a positive correlation between x_t and aggregate output—the familiar Phillips curve trade-off.

Let us trace the sort of adjustment path the model generates following a once and for all increase in m engineered by the monetary authority but underestimated by all the younger generation agents. In the first period following the increase, say period t, the amount of money which must be absorbed by each age-0 agent increases in each market. This will require the current price of real output to increase. But the more unexpected the monetary expansion (the larger x_t and μ_t) the more such a rise in price will be interpreted as a favourable improvement in the intertemporal terms of trade. Age-0 agents will increase output to acquire higher real balances so the market clearing price will rise proportionally less than m. In the next period, when the current younger generation come to spend their money they will be disappointed on average with "what their pounds will buy." But it will be too late to rectify the mistake of the previous period. They may well view the expansion of the previous period as a "good" time of boom and prosperity whereas their disappointment over the shrunken purchasing power

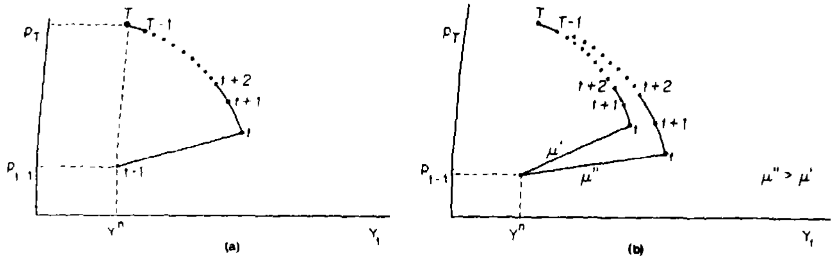


FIG. 2. Learning adjustment paths.

of personal savings will be directed simply at the current inflationary conditions.[12]

The new younger generation in period $t+1$ will not tend to underestimate the monetary expansion to the same degree as their predecessors. For they will have inherited their predecessors' posterior distribution on m. In particular $m_{t+1} > m_t$ because the high price observed in period t will lead to an upward revision.[13] The result will be a *decline* in aggregate output plus a further *rise* in its average price. This process will continue over however many periods are required for the subjective distribution to converge on the new value of m. At that point, on the completion of learning, a new full information equilibrium will be established. Of course, at this new equilibrium, output will be back to the natural rate and the price level will have risen proportionally to the increased money in circulation. The path generated by this sequence of monetary learning equilibria is sketched in Fig. 2(a). The new equilibrium is reached at time T and price p_T.

Consider the effect on the adjustment path of differences in the extent to which agents believe they know the true value of m. One can imagine the case of agents being aware that a monetary expansion has taken place and feeling very uncertain as to its magnitude. The formal equivalent is a low value of the prior precision. On the other hand, agents may feel extremely confident in gauging the extent of the expansion but nevertheless underestimate it or are perhaps even totally unaware, initially, that it has taken place. This corresponds to a high value of the prior precision. The different possible values of μ_t define a family of adjustment curves of which two are shown in Fig. 2(b). From the result of Theorem 5, it is clear that, for the case $m_t < m$, the higher μ_t the further to the right the path will lie.

The interpretation is straightforward: if the authorities can engineer a monetary expansion, without disturbing agents' confidence that the money supply is lower than it really is (or has become) then a more effective tradeoff ensues and it will take longer for the path to return to the natural rate of output. The latter result can be understood directly in terms of the Bayesian revision process. When the posterior mean is calculated as a weighted average of the prior mean and a sample statistic (usually the sample mean) the weight given to the former increases with μ_t. Consequently more weight is given to prior beliefs and less to sample evidence. Hence, if prior beliefs are "wrong," the more observations that are required for the subjective density to converge to the true parameter value.

---

[12] See Lucas [12] for his comments about how agents of both generations will tend to view periods of real expansion followed by inflation.

[13] In Bayesian analysis the posterior mean is usually some weighted average (depending on the precision of the prior p.d.f.) of the prior mean and the sample mean. Hence a high observed price will "pull up" the mean. See DeGroot [5].

It is possible to analyse other case in a similar fashion: an increase in m which is overestimated and, more alien to recent experience, decreases in m which are either over or underestimated. As well as the Phillips curve tradeoff, or lack of, it is possible to see counterparts in the model solutions of such phenomena as stagflation, the apparent variability in learning speeds of agents in different historical periods and the chagrin of policy makers that their expansionary policies would work if only supply behaviour were not so sluggish.

An Illustrative Example

As an example of the model consider the following special case. In fact this example fails to conform to some of our assumptions such as the continuity of the random variables. But this does not matter since the purpose here is strictly to give some tangibility to the somewhat abstract character of model, with as simple an example as possible.

Let the random variable which determines the distribution of the younger generation across the two markets, θ, have just two possible values denoted θ_1 and θ_2, where $P(\theta=\theta_1)=a=1-P(\theta=\theta_2)$, and let the uncertainty of each age-0 agent regarding m consist simply in his not knowing which of two possible values, m_1 or m_2, is the true one. Denote his subjective probability that $m=m_1$ at the beginning of period t by β_t.

Assume that the utility function $U(c_t, n_t)+V(c_{t+1})$ gives rise to a demand and a supply function that are linear in the expected intertemporal relative price $E[p_{t+1} \mid p_t]/p_t$. Specifically suppose

$$c_t = a_1 + a_2 E[p_{t+1} \mid p_t]/p_t, \qquad a_2 > 0, \tag{23}$$

$$n_t = a_3 - a_4 E[p_{t+1} \mid p_t]/p_t, \qquad a_4 > 0,\ a_3 > a_1, \tag{24}$$

are the functions describing optimal choices for period t consumption and labour supply.

The budget constraint (4) with (23) and (24) give the demand for money function:

$$y_t = d_1 p_t - d_2 E[p_{t+1} \mid p_t], \tag{25}$$

where $d_1 = a_3 - a_1$ and $d_2 = a_2 + a_4$; $d_1, d_2 > 0$.

The equilibrium condition equating money demand and supply in period t is

$$\frac{m}{\theta_t} = d_1 p_t - d_2 E[p_t + 1 \mid p_t]. \tag{26}$$

We now look for an equilibrium price function which possesses the SFE property and satisfies (26) for all combinations of the outcomes of the random variables m and θ, i.e., (m_1, θ_1), (m_2, θ_2), (m_1, θ_2) and (m_2, θ_1). Recall that equilibrium price in general reveals only the ratio m/θ. Therefore, to include this essential feature here, as simply as possible, we assume m_1, m_2, θ_1 and θ_2 are such that:

$$\frac{m_1}{\theta_2} < \frac{m_1}{\theta_1} = \frac{m_2}{\theta_2} < \frac{m_2}{\theta_1}. \tag{27}$$

For easy reference denote m_1/θ_2 by y_L, $m_1/\theta_1 = m_2/\theta_1$ by y_M and m_2/θ_1 by y_H.

On the one hand, if the market clears such that $y_t = y_L$ or $y_t = y_H$, the true m becomes known immediately. On the other hand, if $y_t = y_M$ the true m is not revealed. However, the observation enables an agent to revise his subjective probability $P(m = m_1) = \beta_t$ to β_{t+1} in a Bayesian fashion:

$$\begin{aligned} P(m_1 \mid y_M) &\propto P(y_M \mid m_1) P(m_1) \\ &= P(\theta = \theta_1) P(m_1) \\ &= \alpha\beta_t. \end{aligned}$$

Similarly, $P(m_2 \mid y_M) \propto (1 - \alpha)(1 - \beta_t)$.

Hence,

$$\begin{aligned} \beta_{t+1} &= 1 && \text{if } y_L \text{ is observed} \\ &= 0 && \text{if } y_H \text{ is observed} \\ &= \frac{\alpha\beta_t}{\alpha\beta_t + (1 - \alpha)(1 - \beta_t)} && \text{if } y_M \text{ is observed.} \end{aligned}$$

First we look for the long-run equilibrium price function when m is known with certainty. Denote the long-run equilibrium price p_j^i when m_i is known to be true and θ_j is the realized value of θ. The equilibrium conditions which determine p_j^1 are, from (26)

$$\begin{aligned} \frac{m_1}{\theta_1} &= d_1 p_1^1 - d_2(\alpha p_1^1 + (1-\alpha) p_2^1), \\ \frac{m_1}{\theta_2} &= d_1 p_2^1 - d_2(\alpha p_1^1 + (1-\alpha) p_2^1). \end{aligned} \tag{28}$$

A similar pair of linear equations in p_1^2 and p_2^2 determine equilibrium prices when $m = m_2$ with certainty. According to our model, the solutions ought to be expressible in the functional form $p_j^i = m_i \phi(\theta_j)$. Solving the two sets of equations, with a little manipulation, indeed gives

$$p_j^i = \frac{m_i/\theta_j}{d_1 - d_2}\left(d_1 + d_2\left\{\frac{(\theta_j - \theta_2)(1-\alpha)}{\theta_2} + \frac{(\theta_j - \theta_1)\alpha}{\theta_1}\right\}\right). \tag{29}$$

If y_M has been observed in each period from 0 to t, m is still not known with certainty and there exists a price function with the current state of knowledge about m, β_t, as its argument.[14] Denote it $p(\beta_t)$. The possible realization of the random variables in period $t+1$ given y_M has been observed from 0 to t are: (i) $m_1/\theta_2 = y_L$ which reveals m_1 so $p_{t+1} = p_2^1$. The probability of this outcome is $(1-\alpha)\beta_{t+1}$; (ii) $m_2/\theta_1 = y_H$ which reveals m_2, so $p_{t+1} = p_1^2$ with probability $\alpha(1-\beta_{t+1})$; (iii) $m_1/\theta_1 = m_2/\theta_2 = y_M$ so $p_{t+1} = p(\beta_{t+1})$ with probability $\alpha\beta_{t+1} + (1-\alpha)(1-\beta_{t+1})$. Now a momentary equilibrium price function is a price function $p^*(\beta_t)$ which satisfies the market clearing condition:

$$\begin{aligned} y_M = d_1 p^*(\beta_t) - d_2(&\beta_{t+1}\{(1-\alpha) p_2^1 + \alpha p^*(\beta_{t+1})\} \\ &+ (1-\beta_{t+1})\{\alpha p_1^2 + (1-\alpha) p^*(\beta_{t+1})\}) \qquad \text{for all } \beta_t \in (0, 1). \end{aligned} \tag{30}$$

Rearranging, we get

$$\begin{aligned} p^*(\beta_t) = \frac{y_M}{d_1} - \frac{d_2}{d_1}(&\beta_{t+1}\{(1-\alpha) p_2^1 + \alpha p^*(\beta_{t+1})\} \\ &+ (1-\beta_{t+1})\{\alpha p_1^2 + (1-\alpha) p^*(\beta_{t+1})\}), \end{aligned} \tag{31}$$

which is a degenerate integral equation. Unfortunately there does not seem to be a straightforward analytical method of solution.[15] However, consider the mapping defined by

---

[14] In general the variable m/θ would also be an argument of the price function. But since m/θ has only one relevant value here, namely, y_M, for brevity we have not included it explicitly.

$$Tp(\beta_t) \equiv \frac{y_M}{d_1} - \frac{d_2}{d_1}[\beta_{t+1}\{(1-\alpha)p_2^1 + \alpha p(\beta_{t+1})\} + (1-\beta_{t+1})\{\alpha p_1^2 + (1-\alpha)p(\beta_{t+1})\}]. \tag{32}$$

Clearly (31) can be written $p^* = Tp^*$ and since T can be shown to be a contraction mapping if $d_2 < d_1$, by the Banach fixed point theorem, (31) has a unique solution if $d_2 < d_1$. Moreover the solution $p^*(\beta)$ is the limit of the iterative sequence $(p_0, Tp_0, T^2p_0 \cdots)$, where p_0 is any continuous function on $[0, 1]$. Thus we have a method for generating a numerical solution.

There is one special case when it is easy to obtain an analytical solution. When $P(\theta = \theta_1) = \alpha = \frac{1}{2}$ then, if y_M is observed, $\beta_{t+1} = P(m_1 \mid y_M) = \beta_t$. Dropping the * and the time subscript to simplify notation, we solve for $p(\beta)$ from (31) obtaining

$$p(\beta) = \frac{2y_M + d_2(\beta p_2^1 + (1-\beta)p_1^2)}{2d_1 - d_2}$$

and

$$p'(\beta) = \frac{d_2(p_2^1 - p_1^2)}{2d_1 - d_2}.$$

If the condition $d_2 < d_1$ holds, since $p_1^2 > p_2^1$, it follows that $p'(\beta) < 0$. The higher the subjective probability that the lower value of m is true, the lower is the equilibrium price, i.e., if their real prospects appear to improve, age –0 agents will demand the available cash at a lower price with a larger per capita output.

7. CONCLUSIONS

Despite wide acceptance of the natural rate story amongst economists and policy makers, its theoretical underpinnings are surprisingly vague and sketchy. Granted the rational expectations approach has given us plenty of models of what things look like when agents are fully informed and correctly anticipate the future. These are models of the long run, when learning has ceased. Yet virtually nobody claims that such blissful states of full information equilibrium are often attained or, if so, are maintained for any duration. Rather, continuous learning and adjustment to what has been most recently learned are the common characteristics of economic reality. But there are very few models which explicitly describe such learning and adjustment processes. How agents actually acquire their knowledge of the true structure of the economy is an embarrassing question for many rational expectations theorists. Yet the precise adjustment path back to full information following a monetary disturbance, which is a crucial concern of short run macro-economic analysis, depends on the answer. Our model is proposed as an initial approach to providing one.

---

[15] David Kreps has suggested using martingale convergence theorems to obtain a solution This approach may yet bear the required fruit.

In many respects the model we have developed is extremely simple and unrealistic. This is certainly recognised and the results should be treated with appropriate caution. In a more realistic world the personal model of each agent would be correspondingly more complex. In principle, there is no difficulty in extending the momentary equilibrium price function to include all relevant unknown parameters and agents' beliefs about them. For example, an agent may be aware that a group of agents not including himself hold a view about *m* which is different from his own, but in exactly what respect he is at least partially ignorant. Then, according to the present approach, he will form beliefs about their beliefs and he will revise his beliefs conditionally with respect to any relevant signals he may receive.

We have seen how difficult it is to solve for the momentary equilibrium price function even in very simple models. So the assumption that agents can solve out the function in models that come anywhere near to being useful representations of reality is extremely strong. The limited computational ability of agents is almost certain to thwart the attainment of learning equilibrium in the way we have described. One must conclude that to model the formation and modification of expectations realistically, it is essential to take explicit account of the costs and benefits of acquiring *and processing* information. Almost all the rational expectations literature ignores this important problem.[16]

Other extensions of our model are easy to imagine. For example, it would be interesting to investigate the learning adjustment paths when different groups of agents have different access to information or structurally different personal models. In the latter situation, clearly all but possibly one of the personal models must be incorrect specifications of economic reality. It would be important to know under what conditions such erroneous specifications would be revealed by the signals that the agents concerned receive. It is not at all apparent that the available signals will always or even usually do this and thereby lead to corrective respecification. One can imagine agents getting locked-in to an inefficient Nash-type equilibrium in which observed prices or whatever "confirm" their incorrect personal models.[17] This is another problem that is rarely looked in the eye by the natural rate-cum-rational expectations school.

---

[16] Exceptions are the pertinent critiques of rational expectations theorizing by De Canio [4], Friedman [7] and Buiter [2]. On balance we think the experience of developing this model reinforces the scepticism expressed by these writers.

ACKNOWLEDGMENTS

I am grateful for helpful comments from those in attendance when I presented earlier versions of this paper to seminars at the University of Cambridge and the University of Bristol. I am particularly grateful to Angus Deaton and Frank Hahn for advice and encouragement and to an anonymous referee for a number of helpful suggestions. All errors are, of course, my own.

REFERENCES

1. K. J. ARROW AND J. GREEN, "Notes on Expectations Equilibria in Bayesian Settings," Working Paper No. 33, The Economics Series, Institute for Mathematical Studies in the Social Sciences, Stanford, 1973.
2. W. H. BUITER, "The Macroeconomics of Dr. Pangloss: A Critical Survey of the New Macroeconomics," mimeo, University of Bristol, 1979.
3. R. M. CYERT AND M. H. DEGROOT, Rational expectations and Bayesian analysis, *J. Pol. Econ.* **82** (1974).
4. S. J. DECANIO, Rational expectations and learning from experience, *Quart. J. Econ.* **83** (1979).
5. M. H. DEGROOT, "Optimal Statistical Decisions," McGraw–Hill, New York, 1970.
6. P. DIAMOND AND J. E. STIGLITZ, Increases in risk and in risk aversion, *J. Econ. Theory* **8** (1974).
7. B. M. FRIEDMAN, Optimal expectations and the extreme information assumptions of 'rational expectations' macromodels, *J. Monetary Econ.* **5** (1979).
8. S. J. GROSSMAN, Rational expectations and the econometric modelling of markets subject to uncertainty, *J. Econometrics* **3** (1975).
9. S. J. GROSSMAN, Equilibrium under uncertainty and Bayesian adaptive control theory, *in* "Adaptive Economic Models" (R. H. Day and T. Groves, Eds.), Academic Press, New York, 1975.
10. F. H. HAHN, "On the Notion of Equilibrium in Economics," Cambridge Univ. Press, Cambridge, 1973.
11. A. P. KIRMAN, Learning by firms about demand conditions, *in* "Adaptive Economic Models" (R. H. Day and T. Groves, Eds.), Academic Press, New York, 1975.
12. R. E. LUCAS, Expectations and the neutrality of money, *J. Econ Theory* **4** (1972).
13. R. RADNER, Problems in the theory of markets under uncertainty, *Amer. Econ. Rev.* **60** (1970).
14. M. ROTHSCHILD AND J. E. STIGLITZ, Increasing risk *I*: A definition, *J. Econ. Theory* **2** (1970).

---

[17] Arrow and Green [1] and Kirman [11] construct models in which this sort of outcome happens.

JOURNAL OF ECONOMIC THEORY **24**, 265–295 (1981)

Stationary Equilibrium*

TRUMAN F. BEWLEY

Department of Economics, Northwestern University, Evanston, Illinois 60201

Received September 17, 1979; revised April 7, 1980

1. INTRODUCTION

This paper explores a version of general equilibrium theory, which follows naturally from what I term the permanent income hypothesis. By the permanent income hypothesis, I mean the assumption that the marginal utility of money of each consumer is constant. This assumption is an exaggeration of Milton Friedman's permanent income hypothesis [14]. I apply the term stationary equilibrium to the notion of equilibrium suggested by the permanent income hypothesis. A stationary equilibrium is a stationary stochastic process of temporary equilibria. Stationary equilibria are Pareto optimal, even though there are no forward markets for contingent claims. Because of this fact, the stationary equilibrium model may serve as an alternative to the Arrow–Debreu model of equilibrium with contingent claims contracts. However, the stationary equilibrium model should be applied only in the limited setting in which the permanent income hypothesis is appropriate.

The permanent income hypothesis is appropriate only in short-run contexts and when consumer's expectations about the future are fairly stable. In these circumstances, the hypothesis does make sense as a model of consumer behavior which would be accurate up to a close approximation. In order to be convinced that this is so, it should be sufficient to have an intuitive understanding of the nature of the permanent income hypothesis.

One should visualize the consumer as having a nearly fixed idea of the utility of money to him. He buys a unit of a commodity if and only if the utility of the unit bought exceeds the utility of the money given up. Because be behaves in this way, his expenditures fluctuate, and the fluctuations do not necessarily match fluctuations in this income. He uses money to finance the fluctuations in his net expenditures. His utility for money is entirely indirect and may be thought of as a control variable. He determines his

---

* The research reported here was done at the University of Bonn and at Northwestern University, with the support of National Science Foundation Grant SOC-79-06672.

0022-0531/81/020265-31$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

utility of money so that his long-run average expenditure per period equals his long-run average income.

Of course, if the consumer kept his utility of money exactly constant, then his money balances would behave like a random walk and eventually become negative. In order to keep them positive, he must increase his utility of money when the money balances decline. Imagine that the fluctuations in net expenditure were rapid relative to the time horizon and rate of time preference of the consumer. It should be apparent that the consumer would maintain money balances which were large relative to the magnitude of the fluctuations in net expenditure. The order of magnitude of the money balances would then move only sluggishly, so that the same would be true of the consumer's utility for money.

It may be seen that the permanent income hypothesis is appropriate when applied to fluctuations of an everyday nature and when the consumer's economic life is in a state of overall stability.

I have given a rigorous defense of the permanent income hypothesis in another paper [6]. There I prove, roughly speaking, that a consumer's marginal utility of money asymptotically becomes constant as the rapidity of the fluctuations in his environment approaches infinity.

I now describe briefly the model of stationary equilibrium. Consumers' utility functions and endowments and firms' technologies all fluctuate according to a stationary probability law. All trading is for current delivery. Firms act so as to maximize expected long-run average profit per period. The fundamental assumption is that consumers obey the permanent income hypothesis. This amounts to assuming that consumers maximize their expected long-run average flow of utility, subject to the constraint that long-run average expenditure per period not exceed long-run average income. All *firms and consumers have rational expectations. Each firm and consumer* observes the underlying stohastic process which governs all exogenous fluctuations in the economy. A result is that an equilibrium exists in which prices form a stationary stochastic process. The equilibrium defined is called stationary equilibrium.

The assumption that each consumer and firm observes the exogenous stochastic process is quite strong. In fact, consumers and firms do not necessarily need so much information. This issue is discussed in Section 5.

Stationary equilibrium turns out to be equivalent to general equilibrium in a two period economy in which goods in the second period are artifically transferred back to the first period. (This fact is proved in Section 7.) I prove that stationary equilibria exist by proving that the two period economy has an equilibrium. The two period economy necessarily has infinitely many commodities. For prices and allocations depend on the infinite history of past values of the exogenous stationary random process. If this process is not periodic, it has a continuum of possible histories. I prove that the two period *economy has an equilibrium by applying results of a previous paper* [5] on economies with infinitely many commodities.

A routine argument proves that stationary equilibrium is Pareto optimal. It is intuitively fairly clear why markets for contingent claims are not needed for Pareto optimality. In the first place, since the marginal utility of money is constant, a consumer has no need for insurance. In effect, he uses money to insure himself. In the second place, forward markets are not needed to coordinate intertemporal supply and demand, for agents have rational expectations and full information.

Stationary equilibrium may be interpreted as providing a theory of socially optimal inventory policy. In the stationary equilibrium model, a distinction is made between producible and primary goods. Since all fluctuations are thought of as short-lived, one must think of capital goods as primary goods in fixed supply. The producible goods are consumables and intermediate goods. Stocks of producible goods may fluctuate. The quantities held are Pareto optimal since the stationary equilibrium itself is Pareto optimal.

Some caution must be used in interpreting this theory of inventory holdings. Since markets are perfect, some of the usual motives for holding inventories are absent. Agents believe they can always buy whatever they need when they need it.

It is remarkable that Pareto optimal inventory and production decisions are made in stationary equilibrium, even though no individual knows the aggregate quantities held of the various goods. This fact shows how strong are the assumptions of rational expectations and full information.

Stationary equilibrium offers a limited answer to the question of why we do not in reality observe complete markets for contingent claims. Arrow posed this question in [3]. The answer is that self-insurance and rational expectations can take care of every day fluctuations.

If there is one consumer and one firm, then a stationary equilibrium is a stationary optimal consumption and production plan, or a stochastic golden rule. Thus, I give a proof using equilibrium theory of the existence of such a golden rule and associated competitive prices. This problem has been studied by Radner [25], Evstigneev [13], and others. Their work is discussed in Section 6. Of course, the interpretation of stationary equilibrium is very different from that of the golden rule. The golden rule has to do with the theory of optimal capital investment. In stationary equilibrium, capital goods are in fixed supply.

In summary, stationary equilibrium provides a synthesis of general equilibrium theory, the permanent income hypothesis, rational expectations, temporary equilibrium theory and the theory of the golden rule.

Stationary equilibrium is a somewhat technical subject for it requires an infinite dimensional commodity space. Notation and definitions are introduced in the next section. Assumptions are listed in Section 3, and the theorems are listed in Section 4. In Section 5, I discuss certain assumptions. Related work is discussed in Section 6. The proofs follow. These are in large part routine, given knowledge of Debreu [10] and my own paper [5].

2. DEFINITIONS, NOTATION AND THE MODEL

The Underlying Stochastic Process

The exogenous stochastic process is denoted by $\{s_n\}_{n=-\infty}^{\infty}$. It influences utility functions, endowments, and production functions. (Doob [11] is the reference used for stochastic processes.) The random variables, s_n, take values in some measurable space $(M, \mathscr{M})$, where M is a set and $\mathscr{M}$ is a σ-field of subsets of M. The sample space of the process is $S = \{(..., s_{-1}, s_0, s_1, ...) \mid s_n \in M$, for all $n\}$. s denotes an element of S and s_n denotes the nth component of s. The set of all measurable subsets of S is denoted by $\mathscr{S}$, and P denotes the probability on $\mathscr{S}$. $\mathscr{S}$ is the smallest complete σ-field such that all the random variables s_n are measurable with respect to $\mathscr{S}$. Sets $A \in \mathrm{S}$ such that $P(A) = 0$ are called sets of probability zero. An event occurs *almost surely* or *for almost every s* if it occurs for every s except for points s belonging to a set of probability zero.

E denotes the expectation operator corresponding to P. That is, if $X: S \to (-\infty, \infty)$ is integrable with respect to P, then $EX = \int X(s)\, P(ds)$.

$\mathscr{S}_n$ denotes the smallest complete σ-field with respect to which the random variables s_k, $k \leqslant n$, are measurable. Clearly, $\mathscr{S}_n \subset \mathscr{S}$. $\mathscr{S}_n$ represents the information available in period n to an observer of the process.

Let $\sigma: S \to S$ be the shift operator, defined by the formula $(\sigma s)_n = s_{n+1}$. $\{s_n\}$ is *stationary* if and only if σ is probability preserving. That is, σ must satisfy $P(\sigma B) = P(B)$, for every $B \in \mathscr{S}$. A measurable set B is said to be invariant if $P(B) = P(B \cap \sigma B)$. A stationary process is said to be *metrically transitive* if every invariant set is of probability 0 or 1. Metric transitivity is often referred to in probability theory as ergodicity.

I will assume that the process $\{s_n\}$ is *stationary* and *metrically transitive.*

Commodities

There are L types of commodities. $L_c \subset \{1,..., L\}$ denotes the set of consumption goods. $L_0 \subset \{1,..., L\}$ denotes the set of primary commodities, such as labor and raw materials. Capital equipment should also be thought of as a primary good. $L_p = \{k = 1,..., L \mid k \notin L_0\}$ denotes the set of producible goods. Goods not in either L_c or L_0 should be thought of as intermediate goods or goods in process.

Vector Space Notation

R^L denotes L-dimensional Euclidean space. R^{L_c} denotes the subspace of R^L corresponding to L_c. That is, $R^{L_c} = \{x \in R^L \mid x_k = 0 \text{ if } k \notin L_c\}$. R^{L_0} and R^{L_p} are defined similarly.

$\mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P)$ denotes the space of equivalence classes of functions from S to R^L which are measurable with respect to $\mathscr{S}_n$ and are essentially bounded. $f\colon S \to R^L$ is said to be essentially bounded if for some $r > 0$, $P\{s \mid \|f(s)\| > r\} = 0$. The equivalence relation is that of almost sure equality. That is, if $f = S \to R^L$ and $g\colon S \to R^L$ are $\mathscr{S}_n$-measurable and essentially bounded, then f and g represent the same element of $\mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P)$ if $f(s) = g(s)$ for almost every s.

$\mathscr{L}_{1,L}(S, \mathscr{S}_n, P)$ denotes the space of equivalence classes of functions from S to R^L which are $\mathscr{S}_n$-measurable and integrable with respect to P.

$\mathscr{L}_{\infty,L_c}(S, \mathscr{S}_n, P)$, $\mathscr{L}_{\infty,L_p}(S, \mathscr{S}_n, P)$ and $\mathscr{L}_{\infty,L_0}(S, \mathscr{S}_n, P)$ are the subspaces of $\mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P)$ corresponding to L_c, L_p and L_0, respectively.

If $x \in R^L$, then "$x \geqslant 0$" means "$x_k \geqslant 0$, for all k." "$x > 0$" means "$x \geqslant 0$ and $x \neq 0$." "$x \gg 0$" means "$x_k > 0$, for all k." R^L_+ denotes $\{x \in R^L \mid x \geqslant 0\}$, and R^L_- denotes $\{x \in R^L \mid x \leqslant 0\}$. If $x \in R^L$, then $x^+ \in R^L_+$ is defined by $x^+_k = \max(0, x_k)$, for all k.

If $x \in \mathscr{L}_{q,L}(S, \mathscr{S}_n, P)$, where $q = 1$ or ∞, I write "$x \geqslant 0$" if $x(s) \geqslant 0$, for almost every s. I write "$x > 0$" if $x \geqslant 0$ and $x \neq 0$. I write "$x \gg 0$" if $x(s) \gg 0$, for almost every s. Finally, "$x(s) \ggg 0$" means "there exists a positive real number r such that $x_k(s) \geqslant r$ almost surely, for all k."

$\mathscr{L}^+_{\infty,L}(S, \mathscr{S}_n, P)$ denotes $\{x \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P) \mid x \geqslant 0\}$. $\mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_n, P)$, $\mathscr{L}^+_{\infty,L_0}(S, \mathscr{S}_n, P)$ and $\mathscr{L}^+_{1,L}(S, \mathscr{S}_n, P)$ are defined similarly.

If $x \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P)$ and $p \in \mathscr{L}_{1,L}(S, \mathscr{S}_n, P)$, then $p \cdot x$ denotes $\sum_{k=1}^{L} \int p_k(s)\, x_k(s)\, P(ds)$.

If $x \in \mathscr{L}_{q,L}(S, \mathscr{S}_n, P)$, where $q = 1$ or ∞, then $\sigma^m x$ is defined by $\sigma^m x = x(\sigma^m s)$, where σ is the shift operator on S. Clearly, $\sigma^m x \in \mathscr{L}_{q,L}(S, \mathscr{S}_{n+m}, P)$.

Consumers

There are I consumers, where I is a positive integer. The endowment of consumer i is determined by $\omega_i \in \mathscr{L}^+_{\infty,L_0}(S, \mathscr{S}_0, P)$. $\omega_i(\sigma^n s) \in R^{L_0}_+$ is his endowment vector at time n. Note that consumers are endowed only with primary goods.

The utility function of consumer i for consumption in period zero is $u_i\colon R^{L_c}_+ \times S \to (-\infty, \infty)$. I assume that u_i is measurable with respect to $\mathscr{B} \in \mathscr{S}_0$, where $\mathscr{B}$ is the Borel σ-field of $R^{L_c}_+$ and $\mathscr{B} \otimes \mathscr{S}_0$ is the product σ-field generated by $\mathscr{B}$ and $\mathscr{S}_0$. The utility function for period n is $u_i(x, \sigma^n s)$.

Utility is additively separable with respect to time and satisfies the expected utility hypothesis. That is, if $x^n \in \mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_n, P)$, for $n = 1, \dots, N$, then the total utility to consumer i of the bundle $(x^1, \dots, x^n)$ is $\sum_{n=1}^{N} \int u_i(x^n(s), \sigma^n s)\, P(ds)$.

Firms

There are J firms, where J is a positive integer. The production function of firm j is $g_j\colon R^L_- \times S \to R^{L_p}_+$. I assume that g_j is measurable with respect to $\mathscr{B} \otimes \mathscr{S}_1$, where $\mathscr{B}$ is the Borel σ-field on R^L_-. g_j defines the technology at time zero for transforming inputs at time zero into outputs at time one. The technology at time n is defined by $g_j(\cdot\,, \sigma^n s)$. Inputs y_0 at time n may be transformed into outputs y_1 at time $n+1$ when the state of the world is s, if and only if $y_1 \leqslant g_j(y_0, \sigma^n s)$. Inputs carry a negative sign and outputs a positive sign.

Firms are owned by consumers. Consumer i owns a proportion θ_{ij} of firm j, for $i = 1,\ldots, I$ and $j = 1,\ldots, J$. $0 \leqslant \theta_{ij} \leqslant 1$, for all i and j, and $\sum_{i=1}^{I} \theta_{ij} = 1$, for all j.

The Economy

The economy is denoted by $\mathscr{E}$. It is described by the list $\{(u_i, \omega_i), g_j, \theta_{ij}\colon i = 1,\ldots, I \text{ and } j = 1,\ldots, J\}$.

Allocations

A *consumption program* for a particular consumer is of the form $(x^0, x^1,\ldots)$, where $x^n \in \mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_n, P)$, for all n. The consumption vector at time n determined by this program is $x^n(s) \in R^{L_c}_+$.

A *stationary consumption program* is a program such that $x^n = \sigma^n x$, for all n, where $x \in \mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_0, P)$. Such a program is denoted simply by x.

A *production program* is of the form

$$(y_1^{-1}, (y_0^0, y_1^0), (y_0^1, y_1^1),\ldots), \qquad \text{where} \quad y_1^{-1} \in \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_{f0}, P)$$

and

$$(y_0^n, y_1^n) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_{n+1}, P), \qquad \text{for} \quad n \geqslant 0.$$

The program is *feasible* for firm j if $y_0^n \leqslant 0$ and $y_1^n(s) \leqslant g_j(y_0^n(s), s)$ almost surely, for all $n \geqslant 0$. The input vector at time n is $y_0^n(s)$, and the output vector at time $n+1$ is $y_1^n(s)$. Notice that this definition of feasibility implies free disposibility.

A *stationary production program* is a program such that $y_1^{-1} = \sigma^{-1} y_1$ and $(y_0^n, y_1^n) = (\sigma^n y_0, \sigma^n y_1)$, for $n \geqslant 0$, where $(y_{01}, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$. Such a program is denoted simply by (y_0, y_1).

An *allocation* for the economy $\mathscr{E}$ is of the form $((x_i)_{i=1}^{I}, (y_j)_{j=1}^{J})$, where each $x_i = (x_i^0, x_i^1,\ldots)$ is a consumption program and each $y_j = (y_{j1}^{-1}, (y_{j0}^0, y_{j1}^0), (y_{j0}^1, y_{j1}^1),\ldots)$ is a production program feasible for firm j. $((x_i), (y_j))$ is *feasible* if $\sum_{i=1}^{I} x_i^n = \sum_{i=1}^{I} \sigma^n \omega_i + \sum_{j=1}^{J} (y_{j0}^n + y_{j1}^{n-1})$, for all $n \geqslant 0$.

A *stationary allocation* is denoted by $((x_i)_{i=1}^{I}, (y_{j0}, y_{j1})_{j=1}^{J})$, where each x_i

is a stationary consumption program and each (y_{j0}, y_{j1}) is a stationary production program. The allocation is feasible if $\sum_{i=1}^{I} x_i = \sum_{i=1}^{I} \omega_i + \sum_{j=1}^{J} (y_{j0} + \sigma^{-1} y_{j1})$.

The Assumption of Full Information

There is an assumption of full information implied by the definition of allocation. Presumably, agents have enough information to choose the bundles determined by an allocation $((x_i), (y_j))$. At time n, consumer i chooses $x_i^n(s)$, where $x_i = (x_n^0, x_n^1, \ldots)$, and firm j chooses y_{j0}^n, where $y_j = (y_{j1}^{-1}, (y_{j0}^0, y_{j1}^0), \ldots)$. x_i^n, y_{j0}^n, and y_{j1}^{n-1} are all measurable with respect to $\mathscr{S}_n$. That is, they depend on possibly all the values of s_k, for $k \leqslant n$. In order to make such choices, the agents should know s_k, for $k \leqslant n$. That is, they should have full information. In stationary equilibrium, the information requirements are less exaggerated since prices reveal information. I discuss this matter in Section 5.

Pareto Optimality

A feasible allocation $((x_i), (y_j))$ is said to be Pareto optimal if there exists no feasible allocation $((\bar{x}_i), (\bar{y}_j))$ such that

$$\liminf_{N \to \infty} N^{-1} \sum_{n=0}^{N} \int [u_i(\bar{x}_i^n(s), s) - u_i(x_i^n(s), s)] \, P(ds) \geqslant 0, \text{ for all } i,$$

with strict inequality for some i. Observe that Pareto optimality is defined in terms of long-run average flows of expected utility.

Prices

I define only stationary price systems. A *price system* is simply a non-zero element, p, of $L_{1,L}^{+}(S, \mathscr{S}_0, P)$. The vector of prices at time n defined by p is $p(\sigma^n s) \in R_+^L$.

If $x \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P)$, then $p \cdot x$ is the long-term average financial flow determined by p and x. That is, $p \cdot x = \lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} p(\sigma^n s) \cdot x(\sigma^n s)$, almost surely. This fact follows from the strong law of large numbers. The strong law of large numbers applies since $\{s_n\}$ is stationary and metrically transitive. (See Doob [11, p. 465].)

Profit Maximization

Given a stationary price system p, each firm chooses a stationary program so as to maximize its long-run average flow of profit. If the price system is p, the average flow of profit from the stationary program (y_0, y_1) is $p \cdot (y_0 + \sigma^{-1} y_1)$.

The set of stationary programs feasible for firm j is $Y_j = \{(y_0, y_1) \in$

$\mathscr{L}_{\infty,L}(S,\mathscr{S}_0,P)\times\mathscr{L}_{\infty,L_p}(S,\mathscr{S}_1,P) \mid y_0\leqslant 0$ and $y_1(s)\leqslant g_j(y_0(s),s)$ almost surely$\}$. Therefore, the problem of firm j is

$$\max\{p\cdot(y_0+\sigma^{-1}y_1) \mid (y_0,y_1)\in Y_j\}.$$

$\eta_j(p)$ denotes the set of solutions to this problem. $\pi_j(p)$ denotes the maximum profit flow. That is, $\pi_j(p)=p\cdot(y_0+\sigma^{-1}y_1)$, where $(y_0,y_1)\in\eta_j(p)$.

It is easy to see that if $(y_0,y_1)\in\eta_j(p)$, then the infinite horizon program $(\sigma^{-1}y_1, (y_0,y_1), (\sigma y_0,\sigma y_1),\ldots)$ solves the problem

$$\max\left\{\limsup_{N\to\infty} N^{-1}\sum_{n=1}^{N}\int p(\sigma^n s)\cdot(y_0^n(s)+y_1^{n-1}(s))\,P(ds) \mid (y_1^{-1},(y_0^0,y_1^0),\ldots) \text{ is a program feasible for firm } j\right\}.$$

Thus, a program in $\eta_j(p)$ maximizes firm j's long-run average flow of expected profit.

Utility Maximization

Given the stationary price system p, consumer i's stationary budget set is $\beta_i(p)=\{x\in\mathscr{L}_{\infty,L_c}(S,\mathscr{S}_0,P) \mid p\cdot x\leqslant p\cdot\omega_i+\sum_{j=1}^{J}\theta_{ij}\pi_j(p)\}$. His maximization problem is

$$\max\left\{\int u_i(x(s),s)\,P(ds) \mid x\in\beta_i(p)\right\}.$$

$\xi_i(p)$ denotes the set of solutions to this problem.

If $x\in\xi_i(p)$ (and u_i satisfies assumptions 3.5 and 3.6 below), then the program $(x,\sigma x,\sigma^2 x,\ldots)$ solves the infinite horizon maximization problem.

$$\max\left\{\limsup_{N\to\infty} N^{-1}\sum_{n=1}^{N}\int u_{in}(x^n(s),s)\,P(ds) \mid x^n\in\mathscr{L}^{+}_{\infty,L_c}(S,\mathscr{S}_n,P), \text{ for all } n, \text{ and } \limsup_{N\to\infty} N^{-1}\sum_{n=1}^{N} p(\sigma^n s)\cdot(x^n(s)-\omega_i(\sigma^n s))\leqslant 0, \text{ almost surely}\right\}.$$

That is, every $x\in\xi_i(p)$ maximizes the long-run average flow of utility, subject to a long-run average budget constraint. This fact is easy to prove. The proof uses Fatou's lemma (Halmos [18, p. 113]) and the following fact. If $x\in\xi_i(p)$, then there exists a positive number λ_i, such that for almost every s, $u_i(z,s)\leqslant u_i(x(s),s)+\lambda_i p(s)\cdot(z-x(s))$, for every $z\in R_{+}^{L_c}$. That is, consumer i's behavior satisfies the *permanent income hypothesis*.

Stationary Equilibrium

A stationary equilibrium for $\mathscr{E}$ consists of $((x_i), (y_{j0}, y_{j1}), p)$, where

2.1. $((x_i), (y_{j0}, y_{j1}))$ is a feasible stationary allocation,

2.2 p is a stationary price system,

2.3. $(y_{j0}, y_{j1}) \in \eta_j(p)$, for all j, and

2.4. $x_i \in \xi_i(p)$, for all i.

A stationary equilibrium *with transfer payments* consists of $((x_i), (y_{j0}, y_{j1}), p)$, where these satisfy conditions 2.1–2.3 above and the following condition.

2.5. For each i, x_i solves the problem

$$\max \left\{ \int u_i(x(s), s)\, P(ds) \mid x \in \mathscr{L}^{+}_{\infty, L_c}(S, \mathscr{S}_n, P) \text{ and } p \cdot x \leqslant p \cdot x_i \right\}.$$

The flow of transfer payments made by consumer i is

$$p \cdot \omega_i + \sum_{j=1}^{J} \theta_{ij} \pi_j(p) - p \cdot x_i.$$

Associated with any stationary equilibrium with transfer payments are marginal utilities of money, $\lambda_1, \ldots, \lambda_I$. λ_i, the marginal utility of money of consumer i, is the multiplier associated with the budget constraint in 2.5.

3. ASSUMPTIONS

I list below all the assumptions I use. Some have already been mentioned.

The Stochastic Process

3.1. $\{s_n\}_{n=-\infty}^{\infty}$ is stationary and metrically transitive.

Non-Triviality

3.2. $L_c \neq \varnothing$, $L_p \neq \varnothing$ and $L_0 \neq \varnothing$. $I \geqslant 1$ and $J \geqslant 1$.

Consumers

3.3. $\omega_i \in \mathscr{L}^{+}_{\infty, L_0}(S, \mathscr{S}_0, P)$, for all i.

3.4. For all i, $u_i \colon R^{L_c}_{+} \times S \to (-\infty, \infty)$ is measurable with respect to $\mathscr{B} \otimes \mathscr{S}_0$, where $\mathscr{B}$ is the Borel σ-field on $R^{L_c}_{+}$.

3.5. For each i, $u_i(x, \cdot\,) \colon S \to (-\infty, \infty)$ is integrable, for each $x \in R^{L_c}_{+}$.

3.6. For each i and very s, $u_i(\cdot, s): R^{L_c}_+ \to (-\infty, \infty)$ is continuous, concave and weakly monotone.

By weakly monotone, I mean that $u_i(x, s) > u_i(y, s)$ whenever $x \gg y$.

Firms

3.7. (Measurability) For each j, $g_j = R^L_- \times S \to R^{L_p}_+$ is measurable with respect to $\mathscr{B} \otimes \mathscr{S}_1$, where $\mathscr{B}$ is the Borel σ-field on R^L_-.

Assumptions 3.8–3.11 below hold for all j and s.

3.8. (Convexity) Each component of $g_j(y_0, s)$ is a continuous concave function of y_0.

3.9. (Free Disposability) Each component of $g_j(y_0, s)$ is monotonely non-increasing with respect to y_0.

That is, if $y_0 \leqslant \bar{y}_0$, then $g_j(y_0, s) \geqslant g_j(\bar{y}_0, s)$.

3.10. (Constant Returns to Scale) $g_j(ty_0, s) = tg_j(y_0, s)$ for all $t \geqslant 0$.

3.11. (Necessity of Primary Inputs) $g_j(y_0, s) = 0$, if $y_{0k} = 0$, for all $k \in L_0$.

Lipschitz Condition

3.12. There exists $K > 0$ such that for all s and j, $g_j(y_0, s)$ is a Lipschitz function of y_0 with constant K.

That is, $\| g_j(y_0, s) - g_j(\bar{y}_0, s)\| \leqslant K \| y_0 - \bar{y}_0 \|$, for all j and s. This is a technical assumption which is used to prove that production sets are closed. (It is used in the proof of Lemma 8.2.) It is intuitively reasonable, but excludes the convenient Cobb–Douglas production function.

Boundedness

The next assumption guarantees that the set of feasible stationary allocations for the economy is bounded.

3.13. There exists $g: R^L_- \to R^{L_p}_+$ such that $g_j(y_0, s) \leqslant g(y_0)$, for all j, y_0 and s, and g satisfies assumptions 3.8–3.11.

Adequacy

The last two assumptions guarantee that no consumer would have a zero income in an equilibrium. I call these assumptions the adequacy assumptions. The need for adequacy assumptions is a well-known problem in equilibrium theory. The adequacy assumptions seem always to be the most awkward part of any equilibrium existence theorem. The theorem of this paper is no exception. Roughly speaking, my assumptions guarantee that any consumer's endowment makes possible a stationary production program for the whole economy which produces a positive amount of every good.

3.14. $\omega_i \gg 0$, for all i.

This means that each consumer is endowed with every primary good. Recall that $Y_j = \{(y_0, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P) | \, y_0 \leqslant 0$ and $y_1(s) \leqslant g_j(y_0(s), s)$ almost surely$\}$.

3.15. There exist $\omega \in \mathscr{L}^+_{\infty,L_0}(S, \mathscr{S}_0, P)$ and

$$(y_{j0}^n, y_{j1}^n) \in Y_j, \qquad \text{for} \quad j = 1,\dots, J \text{ and } n = 1,\dots, N$$

such that

$$\text{(i)} \qquad \sum_{j=1}^{J} (y_{j0}^n + \sigma^{-1} y_{j1}^{n-1}) + \omega \geqslant 0, \qquad \text{for} \quad n = 1,\dots, N,$$

$$\sum_{j=1}^{J} (y_{j0}^1 + \sigma^{-1} y_{j1}^N) + \omega \geqslant 0$$

and

(ii) there exists a positive number r such that

$$\sum_{j=1}^{J} (y_{j0k}^1(s) + \sigma^{-1} y_{j1k}^N(s)) \geqslant r \qquad \text{almost surely, for all } k \in L_p.$$

Strong Monotonicity

In one theorem, I make use of the following strengthening of the monotonicity of utility assumed in 3.6.

3.16. For each i and every s, $u_i(\cdot\,, s) \colon R_+^{L_c} \to (-\infty, \infty)$ is strongly monotone.

By strongly monotone, I mean that $u_i(x, s) > u_i(y, s)$ whenever $x > y$.

4. THEOREMS

I assume that assumptions 3.1–3.15 apply.

4.1. THEOREM. *There exists a stationary equilibrium.*

4.2. THEOREM. *Let* $\lambda_1, \dots, \lambda_I > 0$. *There exists a stationary equilibrium with transfer payments such that for each* i, λ_i *is the marginal utility of consumer* i *in the equilibrium.*

4 3. THEOREM. *The allocation of any stationary equilibrium with transfer payments is Pareto optimal.*

4.4. THEOREM. ***Assume that the strong monotonicity assumption applies to 3.16. Then, any Pareto optimal stationary allocation is the allocation of a stationary equilibrium with transfer payments.***

Theorem 4.2 expresses better the intuitive notion of stationary equilibrium than does Theorem 4.1. Stationary equilibrium is a way of visualizing what happens in a reasonably stable economy over a period of, say, one year. It is essential to this image that the marginal utility of money of each consumer be constant. It does not really matter whether consumers exactly balance their budgets.

The constancy of the marginal utility of money is, of course, an idealization. I think of each consumer as continually, but slowly, adjusting his marginal utility of money in his attempt to control his finances. It is the slowness that is important. The effect of Theorem 4.1 is simply to reassure us that indeed all consumers could simultaneously have just the right marginal utility of money.

5. DISCUSSION OF ASSUMPTIONS

Information

I have already pointed out in Section 2 that I implicitly assume full information. However, a careful analysis of stationary equilibrium shows that agents do not necessarily need an exorbitant amount of information in stationary equilibrium.

Consider a consumer. If his utility function is strictly concave, than one can think of him as needing to know only current prices and the current state of his utility function. For let p be a stationary equilibrium price system and let λ_i be consumer i's marginal utility of money. If $u_i(\cdot, s)$ is a strictly concave function, then there is one one point $x(s) \in R_+^{L_c}$ satisfying $u_i(z, s) \leqslant u_i(x(s), s) + \lambda_i p(s) \cdot (z - x(s))$, for all $z \in R_+^{L_c}$.

$x(s)$ is the demand of consumer i when state s occurs. Clearly, $x(s)$ is completely determined by $u_i(\cdot, s)$, $p(s)$ and λ_i. Of course, λ_i must be such that the consumer satisfies his long-term average budget constraint. If the consumer were to calculate this constraint, he would have to know the entire joint distribution of $u_i(\cdot, s)$ and $p(s)$. However, one can think of the consumer as having found the appropriate level of λ_i by trial and error.

A firm's problem is more difficult, for the firm must predict the future behavior of prices in making decisions. Let p be a stationary equilibrium price system and let $(y_0, y_1) \in \eta_j(p)$. Then for almost every s, $y_0(s)$ solves the problem.

$$\max_{y_0} [p(s) \cdot y_0 + E(\sigma^{-1} p(\cdot) \cdot g_j(y_0, \cdot) \mid \mathscr{S}_0)(s)].$$

In this expression, $E(\cdot \mid \mathscr{S}_0)$ denotes expectation conditioned on $\mathscr{S}_0$. (This concept is defined in Doob [11, pp. 15–20].) Thus, firm j needs to predict its own production function and the prices of its outputs. Presumably, the firm knows the probability distribution of its own production function. The problem is whether it would be able to predict the prices of its outputs. (Its prediction must be as good as that of anyone else in the economy.) This would be nearly the case in the following situation. Imagine that random variation is of two sorts, generalized (e.g., weather) and specific to individual agents (e.g., machine breakdowns). Suppose that all agents have the same information about generalized events and the variation specific to an individual is statistically independent of all other forms of variation. Then, if there were a great many individual firms and consumers, the variation specific to an individual would have a negligible effect on prices. Only information about generalized events would be relevant for price prediction, and this information would be available to everybody.

Malinvaud [22] has made the distinction between variation specific to individuals and generalized variation. He pointed out its relevance to the theory of insurance and to the Pareto optimality of profit maximization.

Separability of Utility

A fundamental assumption is that utility is additively separable with respect to time. This assumption perhaps makes sense in the context of models of growth theory, where time periods are years or generations. It is hard to defend in the context I have in mind. If time periods are short, consumption in one period should affect appetites in immediately following periods. (This point was made to me forcefully by Christian von Weizsächer.) It makes more sense to suppose that the current flow of utility depends on current and past consumption. If this were so, a consumer might prefer a non-stationary program to any stationary one. I have not studied this problem in detail.

Ownership of Capital Equipment

It may seem strange that firms have no endowment, for I assume that capital goods are not producible and so are part of the economy's endowment. These goods are presumably owned by firms and cannot be exchanged, at least not in the short-run. However, it is possible to interpret my model in such a way that particular capital goods are assigned to particular firms.

Suppose that we wish to assign the bundle ω_j of capital goods to firm j, for each j. First of all, it is necessary to label these goods not only by their physical type but by the label j of the firm to which they are assigned. Thus, ω_j would belong to a subspace of the commodity space which pertains only to firm j. The production function of each firm must be so defined that it can use only capital goods with its own label. Finally, the bundle $\theta_{ij}\omega_j$ must be made part of the endowment of consumer i, for each i. That is, the endowment of each firm is assigned to the firm's owners. In stationary equilibrium, each firm would rent its capital equipment from its own stockholders. These rents may be thought of as the profits of the firm in the every day sense of the word. Since I assume constant returns to scale, the stockholders would receive no income from firms other than these rents.

Constant Returns to Scale

The assumption of constant returns to scale has the disturbing consequence that the scale of output of a firm is not determined by prices. However, if capital goods were fixed in the way I just described, then the output of each firm could indeed be completely determined by prices.

Intermediate Goods

Technologies have been defined as if all production processes required only one period. However, it is possible to represent multiperiod production processes simply by appropriate labeling of commodities. For instance, if a wine must age, then a one period old wine would be a different commodity than a two period old wine. One period old wine would be transformed into two period old wine. This observation has often been made in the literature on intertemporal models.

Intermediate goods may be just as specific to a certain firm as are capital goods. Just as in the case of capital goods, this aspect of reality may be modeled by appropriate labeling of commodities and definition of production functions.

Production of Capital Goods

I assume that capital goods are an endowment, even though is reality they are continually being produced and replaced. It is perhaps appropriate to assume that the total stock of capital goods is fixed over a short period of time. It is not realistic to assume that none are produced. One could introduce an exogenous demand for capital goods. These goods would simply disappear, once produced. Their production would be financed by transfer payments from consumers. The stationary equilibrium prices would elicit the production of these goods. All this is easy to do. The only subtlety is that the demand for capital goods must be included in the adequacy assumption.

Another awkward part of my model is the sharp distinction made between capital goods, which are in fixed supply, and other goods used in production. In reality, of course, goods used in production exist in a spectrum of durability. There is no way of avoiding this difficulty, if one is to reatin the simplification offered by the permanent income hypothesis.

Intertemporal Production

My representation of production is extremely simple in that output is entirely determined by inputs and the state of the world. I give firms no possibility of choosing among techniques. They cannot modify their production methods so as to produce more or less of various types of goods, given a fixed input. Moreover at time n, a firm is not able to trade off more output in some states s_{n+1} against less in other states, unless this trade off is affected by the choice of inputs. (In [25], Radner gives firms both types of choice.)

My production functions, in fact, represent techniques. It is possible to use my model in order to represent an economy in which firms have many techniques, as long as they have only finitely many. It is sufficient to label the techniques as distinct firms. The artificial firms would share the same stockholders and capital equipment.

Free Disposability

I assume free disposability and make a good deal of use of it. (See, for instance, the proof of Lemma 8.15.) The assumption is contained in 3.9 and in the definition of a feasible production program (Section 2). This assumption is particularly disturbing in a model representing short-run activity. It says that unused goods never get in the way. However, as always in equilibrium theory, free disposability could be dropped if sufficient monotonicity were postulated. Suppose that every good is always useful in consumption or can always be used to produce some good useful in consumption. Then, every good would have a positive price in stationary equilibrium and so would never be disposed of. Thus, free disposability need be assumed only hypothetically in the middle of proofs. For instance, in proving existence of stationary equilibrium, one could use free disposability in order to obtain an equilibrium and then at the end show that the assumption was not necessary.

6. REVIEW OF THE LITERATURE

From a technical point of view, a stationary equilibrium is simply a general equilibrium in the usual sense, except that there are infinitely many commodities. This fact is made clear in the next section. Thus, stationary equilibrium is simply an interpretation of general equilibrium, just as the Arrow–Debreu model of equilibrium with contingent claims is an interpretation of general equilibrium. (The Arrow–Debreu model is described in [2] and in [10, Chap. 7].)

General equilibrium with infinitely many commodities has been studied by

Debreu [9], myself [5] and Stigum [29, 30]. Debreu proved that equilibrium is Pareto optimal and that a Pareto optimum is a quasi-equilibrium. In [5], I proved that if the commodity space is $\mathscr{L}_\infty$, then an equilibrium exists with prices in $\mathscr{L}_1$. Stationary equilibrium may be viewed as a interpretation of the equilibrium in [5]. Stigum proved that equilibria exist and are Pareto optimal, and that a Pareto optimum is a quasi-equilibrium. However, he assumed that there are countably many commodities. In stationary equilibrium there are either finitely many or a continuum of commodities. Therefore, one cannot use Stigum's results to prove the existence of stationary equilibrium.

My work is related to much work on the interpretation and extension of equilibrium theory. In Section 5, I pointed out the relevance of Malinvaud's work [22]. My work is related to work on rational expectations and temporary equilibrium. Stationary equilibrium is, in fact, a form of temporary equilibrium with rational expectations. The first work on equilibrium with rational expectations is by Radner [24]. Grandmont [16] has surveyed the large literature on temporary equilibrium. In temporary equilibrium theory, it is assumed that consumers face a different budget constraint in every period. There is no permanent income hypothesis, so that equilibria are not necessarily Pareto optimal.

In Radner's rational expectations equilibrium in [24], different agents could have different information. One can imagine that those in ignorance could use prices to surmise the information learned by others. (In [24], Radner assumed that agents did not use prices in this way.) If agents do use prices to obtain information, then rational expectations equilibria may not exist, as examples by Green [17] and Kreps [21] show. Radner [27] and Allen [1] prove that equilibria exist generically. The literature on information revealing prices is large. A very incomplete list of references includes Futia [15], Jordan [20] and Radner [26].

The idea of rational expectations has appeared in many branches of economics. Shiller [28] surveys applications in macroeconomics. Muth [23] invented the concept of rational expectations.

There seems to be a prejudice within the economics profession that rational expectations and symmetric information imply optimality. This idea finds expression in growth theory, when optimal programs are characterized by separating prices (as in Dana [8] and Evstigneev [13]). The optimal programs can be intrerpreted as competitive equilibria for an economy with many identical consumers, who have rational expectations. The present paper seems to be the first in which the connection between Pareto optimality and rational expectations is made clear in a model with diverse consumers.

Similarly, the idea that consumers are to some extent self-insuring seems to be very old, though I have never seen it expressed formally.

I have already pointed out that a stationary equilibrium becomes a golden rule program if there are only one consumer and one firm. It is, of course, not at all appropriate to use equilibrium theory in order to study the golden rule, for equilibrium theory is inherently much more difficult than the theory of the golden rule. However, the models used in the theory of stochastic growth models ressemble my own.

The first work on multi-sector stochastic growth models seems to be that of Radner [25]. Dana [8] Evstigneev [13] and Jeanjean [19] have also made contributions. Dana, Evstigneev and Jeanjean all addressed larger and more difficult issues that the existence of golden rule programs and their characterization using prices. They proved that there exists an optimal infinite horizon program, starting from given initial conditions, and that this program converges to an optimal stationary program. Brock and Mirman [7] solved these same problems for a one sector model.

7. THE TWO PERIOD ECONOMY

In this section, I show that a stationary equilibrium for the economy $\mathscr{E}$ is equivalent to a general equilibrium in a two period economy $\mathscr{E}_0$.

$\mathscr{E}_0$ is defined as follows. The commodity space is $\mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$. The consumption set of each consumer is $\mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_0, P) \times \{0\} \subset \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L}(S, \mathscr{S}_1, P)$. The utility function of consumer i, $U_i^0 \colon \mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_0, P) \times \{0\} \to (-\infty, \infty)$, is defined by $U_i^0(x, 0) = \int u_i(x(s), s)\, P(ds)$. The endowment of consumer i is $(\omega_i, 0) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$. The production possibility set of firm j is $Y_j = \{(y_0, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P) \mid y_0 \leqslant 0$ and $y_1(s) \leqslant g_j(y_0(s), s)$ almost surely$\}$. An artificial production set Y_0 is introduced. $Y_0 = \{(y_0, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P) \mid y_0 \leqslant -\sigma^{-1} y_1\}$. Y_0 represents the imaginary process of transferring goods backwards or forwards through time to the same state of the world. I let $\theta_{i0} = I^{-1}$, for all i. In summary, $\mathscr{E}_0 = \{U_i^0, (\omega_i, 0), Y_j, \theta_{ij} \colon i = 1,\ldots, I, j = 0, 1,\ldots, J\}$.

An *allocation* for $\mathscr{E}_0$ consists of $((x_i, 0)_{i=1}^I, (y_{j0}, y_{j1})_{j=0}^J)$, where $(x_i, 0) \in \mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_0, P) \times \{0\}$ and $(y_{j0}, y_{j1}) \in Y_j$, for all i and j. This allocation is *feasible* if $\sum_{i=1}^I (x_i, 0) = \sum_{i=1}^I (\omega_i, 0) + \sum_{j=0}^J (y_{j0}, y_{j1})$.

A *price system* for $\mathscr{E}_0$ consists of $(p_0, p_1) \in \mathscr{L}^+_{1,L}(S, \mathscr{S}_0, P) \times \mathscr{L}^+_{1,L_p}(S, \mathscr{S}_1, P)$, such that $(p_0, p_1) > 0$.

If (p_0, p_1) is a price system for $\mathscr{E}_0$, then the *supply correspondence* of firm j is $\eta_j^0(p_0, p_1) = \{(y_0, y_1) \in Y_j \mid p_0 \cdot y_0 + p_1 \cdot y_1 \geqslant p_0 \cdot \bar{y}_0 + p_1 \cdot \bar{y}_1$, for all $(\bar{y}_0, \bar{y}_1) \in Y_j\}$. $\pi_j^0(p_0, p_1)$ denotes the maximum profit of firm j. That is, $\pi_j^0(p_0, p_1) = p_0 \cdot y_0 + p_1 \cdot y_1$, for $(y_0, y_1) \in \eta_j^0(p_0, p_1)$.

The *budget set* of consumer i is $\beta_i^0(p_0, p_1) = \{(x, 0) \in \mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_0, P) \times \{0\} \mid p_o \cdot x \leqslant p_0 \cdot \omega_i + \sum_{j=0}^J \theta_{ij} \pi_j^0(p_0, p_1)\}$. The *demand correspondence* of consumer i is $\xi_i^0(p_0,p_1) = \{(x,0)\in\beta_i^0(p_0,p_1) \mid U_i^0(x,0) \geqslant U_i^0(\bar{x},0)$, for all $(\bar{x},0)\in\beta_i^0(p_0,p_1)\}$.

An equilibrium for $\mathscr{E}_0$ consists of $((x_i,0)_{i=1}^I, (y_{j0},y_{j1})_{j=0}^J, (p_0,p_1))$, where

- (i) $((x_i,0), (y_{j0},y_{j1}))$ is a feasible allocation for $\mathscr{E}_0$,
- (ii) (p_0,p_1) is a price system for $\mathscr{E}_0$,
- (iii) $(y_{j0},y_{j1})\in\eta_j^0(p_0,p_1)$, for $j=0,\ldots,J$, and
- (iv) $(x_i,0)\in\xi_i^0(p_0,p_1)$, for $i=1,\ldots,I$.

I now show that to every equilibrium for $\mathscr{E}_0$, there corresponds a stationary equilibrium for $\mathscr{E}$. It is easy to show that the reverse relation applies.

Let $((x_i,0)_{i=1}^I, (y_{j0},y_{j1})_{j=0}^J, (p_0,p_1))$ be an equilibrium for $\mathscr{E}_0$. I show that there exists $(\bar{y}_{j0},\bar{y}_{j1})_{j=1}^J$ such that $((x_i)_{i=1}^I, (\bar{y}_{j0},\bar{y}_{j1})_{j=1}^J, p_0)$ is a stationary equilibrium for $\mathscr{E}$.

Since $(y_{00},y_{01})\in Y_0$, it follows that

$$-\sigma^{-1}y_{01} \geqslant y_{00}. \tag{7.1}$$

Since $(-\sigma^{-1}y_{01},y_{01})\in Y_0$ and $0 = p_0\cdot y_{00} + p_1\cdot y_{01} = \max\{p_0\cdot\hat{y}_0 + p\cdot\hat{y}_1 \mid (\hat{y}_0,\hat{y}_1)\in Y_0\}$, it follows from 7.1 that

$$0 = p_0\cdot(-\sigma^{-1}y_{01} - y_{00}). \tag{7.2}$$

Let $(\bar{y}_{00},\bar{y}_{01}) = (-\sigma^{-1}y_{01},y_{01})$. Clearly, $(\bar{y}_{00},\bar{y}_{01})\in\eta_0^0(p_0,p_1)$.

Let $(\bar{y}_{10},\bar{y}_{11}) = (y_{10}+\sigma^{-1}y_{01}+y_{00},y_{11})$. Equation (7.1) implies that $(\bar{y}_{10},\bar{y}_{11})\in Y_1$, since Y_1 satisfies free disposability (assumption 3.9). Equation (7.2) implies that $(\bar{y}_{10},\bar{y}_{11})\in\eta_1^0(p_0,p_1)$.

If $j>1$, let $(\bar{y}_{j0},\bar{y}_{j1}) = (y_{j0},y_{j1})$.

I now show that $((x_i)_{i=1}^I, (\bar{y}_{j0},\bar{y}_{j1})_{j=1}^J)$ is a feasible allocation for $\mathscr{E}_0$. Since $((x_i)_{i=1}^I, (y_{j0},y_{j1})_{j=0}^J)$ is a feasible allocation for $\mathscr{E}_0$, it follows that

$$\sum_{i=1}^I x_i = \sum_{i=1}^I \omega_i + \sum_{j=0}^J y_{j0} \quad\text{and}\quad 0 = \sum_{j=0}^J y_{j1}.$$

From the definition of $(\bar{y}_{j0},\bar{y}_{j1})$, it follows immediately that

$$\sum_{i=1}^I x_i = \sum_{i=1}^I \omega_i + \sum_{j=0}^J \bar{y}_{j0} \quad\text{and}\quad 0 = \sum_{j=0}^J \bar{y}_{j1}. \tag{7.3}$$

Also, by the definition of $(\bar{y}_{00},\bar{y}_{01})$,

$$\bar{y}_{00} = -{}^{-1}\bar{y}_{01}. \tag{7.4}$$

The second equality of (7.3) implies that $-\bar{y}_{01} = \sum_{j=1}^J \bar{y}_{j1}$. Using (7.4) I obtain $\bar{y}_{00}=\sum_{j=1}^{J}\sigma^{-1}\bar{y}_{j1}$. Hence by the first equation of (7.3) $\sum_{i=1}^{I}x_i=\sum_{i=1}^{I}\omega_i+\sum_{j=1}^{J}(\bar{y}_{j0}+\sigma^{-1}\bar{y}_{j1})$. That is, $((x_i)_{i=1}^{I},(\bar{y}_{j0},\bar{y}_{j1})_{j=1}^{J})$ is a feasible allocation for $\mathscr{E}$.

I next show that $p_{1k}=\sigma p_{0k}$, for all $k\in L_p$. Suppose that for some $k\in L_p$, $P\{s\in S\mid p_{1k}(s)<p_{0k}(\sigma s)\}>0$. Let $A=\{s\in S\mid p_{1k}(s)<p_{0k}(\sigma s)\}$ and let $(\hat{y}_0,\hat{y}_1)\in Y_0$ be defined by

$$\begin{aligned}\hat{y}_{1m}(s)&=-1, &&\text{if}\quad s\in A \text{ and } m=k\\ &=0, &&\text{if}\quad s\notin A \text{ or } m\neq k,\end{aligned}$$

and

$$\hat{y}_0=-\sigma^{-1}\hat{y}_1.$$

Then, $p_0\cdot\hat{y}_0+p_1\cdot\hat{y}_1=\int_A(p_{0k}(\sigma s)-p_{1k}(s))\,P(ds)>0$. Since Y_0 is a cone with apex zero and $(\bar{y}_0,\bar{y}_1)\in\eta_0^0(p_0,p_1)$, $0=p_0\cdot\bar{y}_0+p_1\cdot\bar{y}_1\geqslant p_0\cdot\hat{y}_0+p_1\cdot\hat{y}_1$. This contradiction shows that $p_{1k}\geqslant\sigma p_{0k}$, for $k\in L_p$. A similar argument shows that $p_{1k}\leqslant\sigma p_{0k}$, for $k\in L_p$. This proves that $p_{1k}=\sigma p_{0k}$, for $k\in L_p$.

It now follows that $(p_0,p_1)>0$ implies $p_0>0$, so that p_0 is a price system for $\mathscr{E}$. Also, if $(y_0,y_1)\in Y_j$, then $p_0\cdot y_0+p_1\cdot y_1=p_0\cdot(y_0+\sigma^{-1}y_1)$. Hence, $(\bar{y}_{j0},\bar{y}_{j1})\in\eta_j^0(p_0,p_1)$ implies that $(\bar{y}_{j0},\bar{y}_{j1})\in\eta_j(p_0)$, for $j=1,\ldots,J$. Similarly, $(x_i,0)\in\xi_i^0(p_0,p_1)$ implies $x_i\in\xi_i(p_0)$, for all i. I have shown that $((x_i)_{i=1}^{I},\ (y_{j0},y_{j1})_{j=1}^{J},p_0)$ satisfies conditions 2.1–2.4 and so is stationary equilibrium for $\mathscr{E}$.

8. LEMMAS

In this section, I prove several lemmas, which allow me to apply the results of [5] to the economy $\mathscr{E}_0$.

Closedness of Production Sets

Recall that $Y_j=\{(y_0,y_1)\in\mathscr{L}_{\infty,L}(S,\mathscr{S}_0,P)\times\mathscr{L}_{\infty,L_p}(S,\mathscr{S}_1,P)\mid y_0\leqslant 0$ and $y_1(s)\leqslant g_j(y_0(s),s)$ almost surely$\}$, for $j=1,\ldots,J$.

8.1. LEMMA. *For $j=1,\ldots,J$, Y_j is closed in the weak-star topology.*

The weak-star topology is defined as follows. Observe that any element p of $\mathscr{L}_{1,L}(S,\mathscr{S}_n,P)$ is a linear functional on $\mathscr{L}_{\infty,L}(S,\mathscr{S}_n,P)$. The linear functional is defined by $x\mapsto p\cdot x$, for $x\in\mathscr{L}_{\infty,L}(S,\mathscr{S}_n,P)$. The weak-star topology on $\mathscr{L}_{\infty,L}(S,\mathscr{S}_n,P)$ is the weakest topology such that each functional $p\in\mathscr{L}_{1,L}(S,\mathscr{S}_n,P)$ is continuous. The weak-star topology on $\mathscr{L}_{\infty,L}(S,\mathscr{S}_0,P)\times\mathscr{L}_{\infty,L_p}(S,\mathscr{S}_1,P)$ is the product of the weak-star topologies on each of the component spaces.

The proof of Lemma 8.1 depends on the following lemma.

8.2. **Lemma.** *Let $y_0 = S \to R^L_-$ be measurable with respect to $\mathscr{S}_0$. For each $k = 1,\dots, L$, there exists an integrable function $h_k \colon S \to R^L$, which is measurable with respect to $\mathscr{S}_1$ and such that for almost every $s \in S$, $g_{jk}(z, s) \leqslant g_{jk}(y_0(s), s) + h_k(s) \cdot (z - y_0(s))$, for all $z \in R^L_-$.*

Proof. I apply the following measurable choice theorem of Aumann [4]. Let (T, μ) be a σ-finite measure space, let X be a standard measure space and let H be a measurable subset of $T \times X$ whose projection on T is all of T. Then, there is a measurable function $h \colon T \to X$ such that $(t, h(t)) \in H$, for almost all $t \in T$.

I apply this theorem with $T = S$, $X = R^L_-$ and $\mu = P$. The σ-field on S is $\mathscr{S}_1$. Any Euclidean space is a standard measure space, as is required by the theorem (see Aumann [4]). I let $H = \{(s, x) \in S \times R^L_- \mid g_{jk}(z, s) \leqslant g_{jk}(y_0(s), s) + x \cdot (z - y_0(s))$, for all $z \in R^L_-\}$. The projection of H onto S is all of S, for by assumption 3.8, $g_{jk}(z, s)$ is a concave function of z, for all z. In order to apply Aumann's theorem, I must show that H is measurable with respect to $\mathscr{S}_1 \otimes \mathscr{B}$, where $\mathscr{B}$ is the Borel σ-field of R^L_-.

For each $z \in R^L_-$, let $H(z) = \{(s, x) \in S \times R^L_- \mid g_{jk}(z, s) - g_{jk}(y_0(s), s) - x \cdot (z - y_0(s)) \leqslant 0\}$. $H(z)$ is clearly measurable with respect to $\mathscr{S}_1 \otimes \mathscr{B}$, being the inverse image of the interval $(-\infty, 0]$ with respect to an $\mathscr{S}_1 \otimes \mathscr{B}$-measurable function. Also, $H = \bigcap \{H(z) \mid z \in R^L_-$ has rational coordinates$\}$. Therefore, H is measurable.

Thus, I may apply Aumann's selection theorem to obtain $h \colon S \to R^L_-$, which is measurable with respect to $\mathscr{S}_1$ and such that $(s, h(s)) \in H$ almost surely. By the assumption of free disposability (3.9), $h(s) \leqslant 0$ almost surely. Let h_k be defined by $h_{kn}(s) = \min(h_n(s), K)$, for $n = 1,\dots, L$, where K is the Lipschitz constant of assumption 3.12. It follows from this assumption that $(s, h_k(s)) \in H$ almost surely. Clearly, h_k is $\mathscr{S}_1$-measurable. Since it is essentially bounded, it is integrable. Q.E.D.

Proof of Lemma 8.1. Let (y_0^α, y_1^α), $\alpha \in A$, be a net in Y_j or a generalized sequence, where A is a directed set. (See Dunford and Schwartz [12, p. 26], for a definition of generalized sequence.) Suppose that (y_0^α, y_1^α) converges to $(y_0, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$ in the weak-star topology. I must show that for each $k = 1,\dots, L$, $y_{1k}(s) \leqslant g_{jk}(y_0(s), s)$ almost surely. It is sufficient to show that

$$\int_B y_{1k}(s)\, P(ds) \leqslant \int_B g_{jk}(y_0(s), s)\, P(ds), \qquad \text{for every } B \in \mathscr{S}_1. \qquad (8.3)$$

By the previous lemma, there exists $h_k \colon S \to R^L_-$, which is integrable and $\mathscr{S}_1$-measurable and such that for almost every s, $g_{jk}(z, s) \leqslant g_{jk}(y_0(s), s) + h_k(s) \cdot (z - y_0(s))$. Substituting $y_0^\alpha(s)$ for z, I obtain $y_1^\alpha(s) \leqslant g_{jk}(y_0^\alpha(s), s) \leqslant$

$g_{jk}(y_0(s), s) + h_k(s) \cdot (y_0^\alpha(s) - y_0(s))$. Integrating this inequality over B, I have $\int_B y_1^\alpha(s)\, P(ds) \leqslant \int_B g_{jk}(y_0(s), s)\, P(ds) + \int_B h_k(s) \cdot (y_0^\alpha(s) - y_0(s))\, P(ds)$. Since $\lim_\alpha \int_B y_1^\alpha(s)\, P(ds) = \int_B y_1(s)\, P(ds)$, it is sufficient to prove that

$$\lim_\alpha \int_B h_k(s) \cdot (y_0^\alpha(s) - y_0(s))\, P(ds) = 0. \tag{8.4}$$

Since y_0^α converges to y_0 in the weak-star topology of $\mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P)$, (8.4) would be obvious if B belonged to $\mathscr{S}_0$ and h_k were $\mathscr{S}_0$-measurable. But $B \in \mathscr{S}_1$, and h_k is $\mathscr{S}_1$-measurable. Therefore, I must use conditional expectations.

Let $f\colon S \to R^L_-$ be defined by $f(s) = E(\chi_B(\cdot)\, h_k(\cdot) \mid \mathscr{S}_0)(s)$, where χ_B is the indicator function of B defined by

$$\begin{aligned} \chi_B(s) &= 1, \qquad \text{if} \quad s \in B \\ &= 0, \qquad \text{otherwise.} \end{aligned}$$

By the definition of weak-star convergence, $\lim_\alpha \int f(s) \cdot (y_0^\alpha(s) - y_0(s))\, P(ds) = 0$. But $\int f(s) \cdot (y_0^\alpha(s) - y_0(s)) P(ds) = E[E(\chi_B h_k | \mathscr{S}_0) \cdot (y_0^\alpha - y_0)] = E[E(\chi_B h_k \cdot (y_0^\alpha - y_0) | \mathscr{S}_0)] = \int_B h_k(s) \cdot (y_0^\alpha(s) - y_0(s))\, P(ds)$, by the elementary properties of conditional expectations (see Doob [11, pp. 22 and 35]). This proves (8.4). Q.E.D.

Recall that $Y_0 = \{(y_0, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P) \mid y_0 \leqslant -\sigma^{-1} y_1\}$.

8.5. LEMMA. *Y_0 is closed in the weak-star topology.*

Proof. Let (y_0^α, y_1^α), $\alpha \in A$, be a convergent net in Y_0 with $\lim_\alpha (y_0^\alpha, y_1^\alpha) = (y_0, y_1)$. Because σ is measure preserving, it follows that $\sigma^{-1} y_1^\alpha$ converges to $\sigma^{-1} y_1$. Hence, $\lim_\alpha(-y_0^\alpha - \sigma^{-1} y_1^\alpha) = -y_0 - \sigma^{-1} y_1$. Since $-y_0^\alpha - \sigma^{-1} y_0^\alpha \geqslant 0$, for all α, it follows that $-y_0 - \sigma^{-1} y_1 \geqslant 0$. That is, $(y_0, y_1) \in Y_0$. Q.E.D.

Boundedness

The next lemma asserts that the feasible allocations of $\mathscr{E}_0$ are uniformly bounded. The proof of this lemma depends on assumption 3.13. Before stating the lemma, I must define the appropriate norm. If $y \in R^L$, let $|y| = \max_k |y_k|$. If $y \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_n, P)$, let $|y| = \operatorname{ess\,sup} |y(s)| \equiv \inf\{r > 0 : |y(s)| \leqslant r$ almost surely$\}$.

Let Y_j, for $j = 0, 1, \ldots, J$ be as in the definition of $\mathscr{E}_0$.

8.6. LEMMA. *Let $\omega \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P)$ and let $B = \{(y_{j0}, y_{j1})_{j=0}^J \mid (y_{j0}, y_{j1}) \in Y_j$, for all j and $\sum_{j=0}^J (y_{j0}, y_{j1}) + (\omega, 0) \geqslant 0\}$. Then, B is bounded with respect to the norm $|\cdot|$.*

I prove Lemma 8.6 by means of a sequence of lemmas which follow. Let $Y = \{(y_0, y_1) \in R^L \times R^{L_p} \mid y_0 \leqslant 0$ and $y_1 \leqslant g(y_0)\}$, where g is as in assumption 3.13.

8.7. Lemma. *Let* $t_0 > 0$ *and* $t_1 > 0$. *Then,* $A = \{(y_0, y_1) \in Y \mid -t_0 \leqslant y_{0k}$, *for all* $k \in L_0$, $-t_0 \leqslant y_{1k}$, *for all* k, *and* $|y_0| \leqslant |y_1| + t_1\}$ *is bounded.*

Proof. If the lemma were false, there would exist a sequence $(y_0^n, y_1^n) \in A$ such that $\lim_{n \to \infty} |y_1^n| = \infty$. Let $(\bar{y}_0^n, \bar{y}_1^n) = |y_1^n|^{-1}(y_0^n, y_1^n)$. Then, $|(\bar{y}_0^n, \bar{y}_1^n) \leqslant 1 + |y_1^n|^{-1} t_1$. Hence, $(\bar{y}_0^n, \bar{y}_1^n)$ has a convergent subsequence, which I index by n again. Let $(\bar{y}_0, \bar{y}_1)$ be the limit. Since Y_0 is a closed cone with apex zero, $(\bar{y}_0, \bar{y}_1) \in A$. Also, $\bar{y}_{0k} \geqslant 0$, for $k \in L_0$, and $\bar{y}_1 \geqslant 0$ and $|\bar{y}_1| = 1$. Hence, $\bar{y}_1 > 0$ even though there are no primary inputs. This contradicts assumption 3.11, as applied to g. Q.E.D.

Let $\hat{Y} = \{(y_0, y_1) \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P) \mid y_0 \leqslant 0$ and $y_1(s) \leqslant g(y_0(s))$ almost surely$\}$. By assumption 3.13, $Y_j \subset \hat{Y}$, for $j = 1, \ldots, J$.

8.8. Lemma. *Given* $\omega \in \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P)$, *the set* $A = \{(y_0, y_1) \in \hat{Y} \mid y_0 + \sigma^{-1} y_1 + \omega \geqslant 0\}$ *is bounded with respect to the norm* $|\cdot|$.

Proof. By Lemma 8.7 there exists $T > 0$ such that $|y_1| \leqslant T$ if $(y_0, y_1) \in Y$ satisfies the following conditions: $-|\omega| \leqslant y_{0k}$, for all $k \in L_0$, $-|\omega| \leqslant y_{1k}$, for all k, and $|y_0| \leqslant |y_1| + |\omega| + 1$.

Let $(y_0, y_1) \in A$. Then, if $k \in L_0$,

$$-|\omega| \leqslant -\omega_{lk}(s) = -\omega_{lk}(s) - \sigma^{-1} y_{1k}(s) \leqslant y_{0k}(s) \qquad \text{almost surely.} \quad (8.9)$$

If $k \in L_p$, then

$$y_{1k}(s) \geqslant -y_{0k}(\sigma s) - \omega_k(\sigma s) \geqslant -\omega_k(\sigma s) \geqslant -|\omega| \qquad \text{almost surely.} \quad (8.10)$$

Let $\bar{s} \in S$ be such that $|y_1(\bar{s}) \geqslant |y_1| - 1$ and the following are true. $0 \leqslant y_0(\bar{s}) + y_1(\sigma^{-1}\bar{s}) + \omega(\bar{s})$; (8.9) and (8.10) apply to $(y_0(\bar{s}), y_1(\bar{s}))$; $y_0(\bar{s}) \leqslant 0$; and $y_1(\bar{s}) \leqslant g(y_0(\bar{s}))$. The last two inequalities simply assert that

$$(y_0(\bar{s}), y_1(\bar{s})) \in Y. \qquad (8.11)$$

Also, $0 \leqslant -y_1(\bar{s}) \leqslant y_1(\sigma^{-1}\bar{s}) + \omega(\bar{s})$, so that

$$\begin{aligned} |y_0(\bar{s})| &\leqslant |y_1(\sigma^{-1}\bar{s})| + |\omega(\bar{s})| \leqslant |y_1| + |\omega| \\ &\leqslant |y_1(\bar{s})| + |\omega| + 1. \end{aligned} \qquad (8.12)$$

It follows from (8.9)–(8.12) that $|y_1(\bar{s})| \leqslant T$, so that $|y_1| \leqslant T + 1$. Hence, $|y_0| \leqslant |y_1| + |\omega| \leqslant T + |\omega| + 1$. Q.E.D.

Proof of Lemma 8.6. Let $(y_{j0}, y_{j1})_{j=0}^{J} \in B$. It follows easily that $\omega + \sum_{j=1}^{J}(y_{j0} + \sigma^{-1}y_{j1}) \geqslant 0$ (The argument used here is similar to that just following Eq. 7.4.) $(y_{j0}, y_{j1}) \in \hat{Y}$, for $j = 1,\dots, J$, since $Y_j \subset \hat{Y}$, for such j. Since $\hat{Y}$ is a cone, $\sum_{j=1}^{J} (y_{j0}, y_{j1}) \in \hat{Y}$. Hence by Lemma 8.8, $\{\sum_{j=1}^{J} (y_{j0}, y_{j1}) \mid (y_{j0}, y_{j1})_{j=0}^{J} \in B$, for some $(y_{00}, y_{01})\}$ is bounded. Since $y_{j0} \leqslant 0$, for $j = 1,\dots, J$, it follows that $C \equiv \{(y_{j0})_{j=1}^{J} \mid (y_{j0}, y_{j1})_{j=0}^{J} \in B$, for some (y_{00}, y_{01}) and some $(y_{j1})_{j=1}^{J}\}$ is bounded. Let $T_0 = \max\{\max_{j=1,\dots,J} |y_{j0}|: (y_{j0})_{j=1}^{J} \in C\}$.

Observe that $(y_0, y_1) \in Y$ implies that $(y_0, y_1^{+}) \in Y$. Hence by Lemma 8.7 with $t_0 = t_1 = T_0$, I obtain that $D \equiv \{(y_{j1}^{+})_{j=1}^{J} \mid (y_{j0}, y_{j1})_{j=0}^{J} \in B$, for some (y_{00}, y_{01}) and some $(y_{j0})_{j=1}^{J}\}$ is bounded. Let $T_1 = \max\{\max_{j=1,\dots,J} |y_j^{+}|: (y_{j1}^{+})_{j=1}^{J} \in D\}$. Since $\sum_{j=1}^{J} (y_{j0} + \sigma^{-1}y_{j1}) + \omega \geqslant 0$ and $y_{j0} \leqslant 0$, for $j = 1,\dots, J$, it follows that $y_{j1k}(s) \geqslant -(J-1)\, T_1$, for all k. This proves that $\{(y_{j0}, y_{j1})_{j=1}^{J} \mid (y_{j0}, y_{j1})_{j=0}^{J} \in B$, for some $(y_{00}, y_{01})\}$ is bounded.

If $(y_{j0}, y_{j1})_{j=0}^{J} \in B$, then $-\omega - \sum_{j=1}^{J} y_{j0} \leqslant y_{00} \leqslant -\sigma^{-1}y_{01} \leqslant \sum_{j=1}^{J} \sigma^{-1}y_{j1}$. It follows at once that B is bounded. Q.E.D.

Adequancy

The next lemma states that $\mathscr{E}_0$ satisfies what in [5] I called the "adequancy assumption."

8.13. Lemma. *For each i, there exist $(y_{j0}, y_{j1}) \in Y_j$, for $j = 0, 1,\dots J$, such that $\sum_{j=0}^{J} (y_{j0}, y_{j1}) + (\omega_i, 0) \gg 0$.*

The proof of this lemma makes use of the next lemma.

8.14. Lemma. *There exist $\omega \in \mathscr{L}_{\infty,L_0}^{+}(S, \mathscr{S}_0, P)$ and $(y_{j0}, y_{j1}) \in Y_j$, for $j = 1,\dots, J$, such that*

(i) $\sum_{j=1}^{J} (y_{j0} + \sigma^{-1}y_{j1}) + \omega \geqslant 0$ and

(ii) for some $r > 0$, $\sum_{j=1}^{J} (y_{j0k}(s) + \sigma^{-1}y_{j1k}(s)) \geqslant r$ almost surely, for all $k \in L_p$.

Proof. Let r, ω and (y_{j0}^{n}, y_{j1}^{n}), for $j = 1,\dots, J$ and $n = 1,\dots, N$, be as in assumption 3.15. Let $(y_{j0}, y_{j1}) = N^{-1} \sum_{n=1}^{N}(y_{j0}^{n}, y_{j1}^{n})$. Then,

$$\sum_{j=1}^{J} (y_{j0} + \sigma^{-1}y_{j1}) + \omega$$

$$= N^{-1}\left[\sum_{n=2}^{N} \left(\sum_{j=1}^{J} (y_{j0}^{n} + \sigma^{-1}y_{j1}^{n-1}) + \omega\right) + (y_{j0}^{1} + \sigma^{-1}y_{j1}^{N}) + \omega\right] \geqslant 0,$$

by assumption 3.15. Thus, (i) of the lemma is valid. For $k \in L_p$, $\sum_{j=1}^{J} (y_{j0k}(s) + \sigma^{-1}y_{j1k}(s)) \geqslant N^{-1} \sum_{j=1}^{J} (y_{j0k}^{1}(s) + \sigma^{-1}y_{j1k}^{N}(s)) \geqslant N^{-1}r$, which is condition (ii) of the lemma. Q.E.D.

Proof of Lemma 8.13. Let $\bar{r}$, $\bar{\omega}$ and $(\bar{y}_{j0}, \bar{y}_{j1}) \in Y_j$, for $j = 1,\ldots, J$, satisfy the conditions of Lemma 8.14. By assumption 3.14, there exists $t < 0$ such that $t\bar{\omega}_k(s) + t \leqslant \bar{\omega}_{ik}(s)$ almost surely, for all $k \in L_0$. Let $(y_{j0}, y_{j1}) = t(\bar{y}_{j0}, \bar{y}_{j1})$, for $j = 1,\ldots, J$. By assumption 3.10, $(y_{j0}, y_{j1}) \in Y_j$, for all j. Let (y_{00}, y_{01}) be defined by $(y_{00k}(s),\ y_{01k}(s)) = (\sum_{j=1}^{J} \sigma^{-1} y_{j1k}(s) - 2^{-1}t\bar{r}, -\sum_{j=1}^{J} y_{j1k}(s) + 2^{-1}t\bar{r})$, for all s and for $k \in L_p$. Clearly, $(y_{00}, y_{01}) \in Y_0$.

I now show that $\sum_{j=1}^{J} (y_{j0}, y_{j1}) + (\omega_i, 0) \gg 0$. Let $k \in L_p$. Then, $\sum_{j=0}^{J} y_{j1k}(s) = 2^{-1}t\bar{r} > 0$ and $\sum_{j=0}^{J} y_{j0k}(s) \geqslant \sum_{j=1}^{J} (y_{j0k}(s) + \sigma^{-1} y_{j1k}(s)) - 2^{-1}t\bar{r} \geqslant 2^{-1}t\bar{r} > 0$, almost surely. If $k \in L_0$, then $\sum_{j=0}^{J} y_{j0k}(s) + \omega_{ik}(s) = \sum_{j=1}^{J} y_{j0k}(s) + \omega_{ik}(s) \geqslant t(\sum_{j=1}^{J} y_{j0k}(s) + \bar{\omega}_k(s)) + t \geqslant t > 0$, almost surely. This proves that $\sum_{j=0}^{J} (y_{j0}, y_{j1}) + (\omega_i, 0) \gg 0$. Q.E.D.

The Exclusion Assumption

I now show that what I called the "exclusion assumption" in [5] applies to $\mathscr{X}_0$. In order to describe this assumption, I must introduce some new terminology.

$\mathscr{L}_\infty(S, \mathscr{S}_n, P)$ denotes $\mathscr{L}_{\infty,1}(S, \mathscr{S}_n, P)$, $ba(S, \mathscr{S}_n, P)$ denotes the set of bounded additive set functions defined on $\mathscr{S}_n$ which are absolutely continuous with respect to P. $ba(S, \mathscr{S}_n, P)$ is the set of linear functionals on $\mathscr{L}_\infty(S, \mathscr{S}_n, P)$ continuous with respect to the supremum norm $|\cdot|$.

If $\nu \in ba(S, \mathscr{S}_n, P)$ is non-negative, then ν is said to be purely finitely additive if $\varphi = 0$ whenever φ is a countably additive set function defined on $\mathscr{S}_n$ such that $0 \leqslant \varphi \leqslant \nu$. A theorem of Yosida and Hewitt asserts that if $\nu \in ba(S, \mathscr{S}_n, P))$ and $\nu \geqslant 0$, then there exist $\nu_c \geqslant 0$ and $\nu_p \geqslant 0$ such that ν_c is countably additive, ν_p is purely finitely additive and $\nu = \nu_c + \nu_p$ [31, p. 52, Theorem 1.23]). ν_c is termed the countably additive part of ν and ν_p is termed the purely finitely additive part of ν.

Another theorem of Yosida and Hewitt asserts that if $\nu_p \in ba(S, \mathscr{S}_n, P)$ is purely finitely additive and $\nu_c \in ba(S, \mathscr{S}_n, P)$ is countably additive, then there exists a sequence of sets $S_n \in \mathscr{S}_0$ such that $\nu_p(S \backslash S_n) = 0$, for all n, and $\lim_{n\to\infty} \nu_c(S_n) = 0$ [31, p. 50, Theorem 1.19].

$ba_L^+(S, \mathscr{S}_n, P)$ denotes $\{(\nu_k)_{k=1}^{L} \mid \nu_k \in ba(S, \mathscr{S}_n, P)$ and $\nu_k \geqslant 0$, for all $k\}$. If $\nu \in ba_L^+(S, \mathscr{S}_n, P)$, $\nu = \nu_c + \nu_p$ is the Yosida–Hewitt decomposition of ν. If $(\nu_0, \nu_1) \in ba_L^+(S, \mathscr{S}_0, P) \times ba_{L_p}^+(S, \mathscr{S}_1, P)$, then the decomposition $(\nu_0, \nu_1) = (\nu_{0c}, \nu_{1c}) + (\nu_{0p}, \nu_{1p})$ is defined similarly.

If $\nu \in ba_L(S, \mathscr{S}_n, P)$ and $A \in \mathscr{S}_n$, then $\nu(A) = (\nu_k(A))_{k=1}^{L}$.

Exclusion Assumption. $Y \subset \mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$ satisfies the exclusion assumption if for every $(\nu_0, \nu_1) \in ba_L^+(S, \mathscr{S}_0, P) \times ba_{L_p}^+(S, \mathscr{S}_1, P)$, there exist sequences $S_{0n} \in \mathscr{S}_0$ and $S_{1n} \in \mathscr{S}_1$, $n = 1, 2,\ldots,$ such that

(1) $\lim_{n\to\infty} \nu_{0c}(S_{0n}) = \lim_{n\to\infty} \nu_{1c}(S_{1n}) = 0,$

(2) $\nu_{0p}(S\backslash S_{0n}) = \nu_{1p}(S\backslash S_{1n}) = 0$, for all n and

(3) if $(y_0, y_1) \in Y$, then $(y_0\chi_{S\backslash S_{0n}}, y_1\chi_{S\backslash S_{1n}}) \in Y$, for all n.

In (3) above, $\chi_{S\backslash S_{in}}$ is the indicator function of the set $S\backslash S_{in}$, for $i = 0, 1$. Also, $y_i\chi_{S\backslash S_{in}}$ is defined by $(y_i\chi_{S\backslash S_{in}})(s) = \chi_{S\backslash S_{in}}(s)\, y_i(s)$, for $i = 0, 1$.

8.15. Lemma. *The production sets Y_j, $j = 0, 1, \ldots, J$, of the economy $\mathscr{E}_0$ satisfy the exclusion assumption.*

Proof. Let $(\nu_0, \nu_1) \in ba_L^+(S, \mathscr{S}_0, P) \times ba_{L_p}^+(S, \mathscr{S}_1, P)$ and let $(\nu_0, \nu_1) = (\nu_{0c}, \nu_{1c}) + (\nu_{0p}, \nu_{1p})$ be the Yosida–Hewitt decomposition of (ν_0, ν_1).

For each $k = 1, \ldots, L$, there exists a sequence $A_{0kn} \in \mathscr{S}_0$, $n = 1, 2, \ldots$, such that $\nu_{0pk}(S\backslash A_{0kn}) = 0$ and $\lim_{n\to\infty} P(A_{0kn}) = 0$. Let $S_{0n} = \bigcup_{k=1}^{L} A_{0kn}$. Then $\lim_{n\to\infty} P(A_{0kn}) = 0$ and so $\lim_{n\to\infty} \nu_{0c}(S_{0n}) = 0$. Clearly, $\nu_{0p}(S\backslash S_{0n}) = 0$.

For each $k \in L_p$, there exists a sequence $A_{1kn} \in \mathscr{S}_1$, $n = 1, 2, \ldots$, such that $\nu_{1pk}(S\backslash A_{1kn}) = 0$ and $\lim_{n\to\infty} P(A_{1kn}) = 0$.

Now consider Y_j, for $j \geqslant 1$. Let $S_{1n} = S_{0n} \cup (\bigcup_{k\in L_p} A_{1kn})$. Then, $\lim_{n\to\infty} \nu_{1c}(S_{1n}) = 0$ and $\nu_{1p}(S\backslash S_{1n}) = 0$.

I must show that if $(y_0, y_1) \in Y_j$, then $(y_0\chi_{S\backslash S_{0n}}, y_1\chi_{S\backslash S_{1n}}) \in Y_j$, for all n. That is, I must show that $\chi_{S\backslash S_{1n}}(s)\, y_1(s) \leqslant g_j(\chi_{S\backslash S_{0n}}(s)\, y_0(s), s)$, almost surely. If $s \in S_{1n}$, $\chi_{S\backslash S_{1n}}(s)\, y_1(s) = 0 \leqslant g_j(\chi_{S\backslash S_{0n}}(s)\, y_0(s), s)$, since g_j maps into $R_+^{L_p}$. If $s \notin S_{1n}$, then $\chi_{S\backslash S_{1n}}(s)\, y_1(s) = y_1(s) \leqslant g_j(y_0(s), s) = g_j(\chi_{S\backslash S_{0n}}(s)\, y_0(s), s)$. This completes the proof that Y_j satisfies the exclusion assumption if $j \geqslant 1$.

I now show that Y_0 satisfies the exclusion assumption. Let $S_{0n} = (\bigcup_{k=1}^{L} A_{0kn}) \cup (\bigcup_{k\in L_p} \sigma^{-1} A_{1kn})$ and let $S_{1n} = \sigma S_{0n}$. Since P is invariant with respect to σ, it follows that $\lim_{n\to\infty} P(S_{0n}) = \lim_{n\to\infty} P(S_{1n}) = 0$. Hence, $\lim_{n\to\infty} \nu_{0c}(S_{0n}) = \lim_{n\to\infty} \nu_{1c}(S_{1n}) = 0$. Also, $\nu_{0p}(S\backslash S_{0n}) = \nu_{1p}(S\backslash S_{1n}) = 0$, for all n. It should be clear from the definition of Y_0 that $(y_0\chi_{S\backslash S_{0n}}, y_1\chi_{S\backslash S_{1n}}) \in Y_0$, if $(y_0, y_1) \in Y_0$. Hence, Y_0 satisfies the exclusion assumption. Q.E.D.

9. Proof of Theorem 4.1

By what has been shown in Section 7, it is enough to prove that $\mathscr{E}_0$ has an equilibrium. In order to prove that this is so, I apply results from [5].

In [5], the commodity space is $\mathscr{L}_\infty(M, \mathscr{M}, \mu)$, where $(M, \mathscr{M}, \mu)$ is a σ-finite measure space. ($\mathscr{L}_\infty(M, \mathscr{M}, \mu)$ is $\mathscr{L}_{\infty,1}(M, \mathscr{M}, \mu)$, in the notation of this paper.) The first task is to interpret the commodity space of $\mathscr{E}_0$ as such a space. Recall that the commodity space of $\mathscr{E}_0$ is $\mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$. Let $M = S \times (\{a_1, \ldots, a_L\} \cup L_p)$, where $a_1, \ldots, a_L$ are L distinct symbols, none of them included in L_p. The measurable subsets of $S \times \{a_k\}$, for $k = 1, \ldots, L$, are of the form $A \times \{k\}$, where $A \in \mathscr{S}_0$. Similarly, the measurable subsets of $S \times \{k\}$, for $k \in L_p$, are of the form $A \times \{k\}$, where $A \in \mathscr{S}_1$. μ is defined by $\mu(A \times \{a_k\}) = P(A)$ and $\mu(A \times \{k\}) = P(A)$. It should be obvious how to identify $\mathscr{L}_\infty(M, \mathscr{M}, \mu)$ with $\mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$.

A key assumption of [5] is that preference orderings are continuous with respect to the Mackey topology. I do not define this topology here. (It is defined in [5].) It is sufficient to point out that by Appendix II of [5], the utility functions $U_i^0(x, 0) = \int u_i(x(s), s)\, P(ds)$ are Mackey continuous.

Another key assumption of [5] is that production possibility sets are Mackey closed. The weak-star topology is weaker than the Mackey topology. Hence, Lemmas 8.1 and 8.5 imply that the production possibility sets of $\mathscr{E}_0$ are Mackey closed.

In [5], I also assumed that production possibility sets are convex cones with apex zero. This is obviously true of Y_0. That it is true of Y_j, for $j \geqslant 1$, follows from assumptions 3.8 and 3.10.

Another assumption of [5] is the "monotonicity assumption." In order that this assumption be satisfied, it is enough that

(1) $Y_j - (\mathscr{L}^+_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}^+_{\infty,L_p}(S, \mathscr{S}_1, P)) \subset Y_j$, for some j, and

(2) $L_c \neq \varnothing$, and $U_i(x, 0) > U_i(z, 0)$, if $x \gg z$ and x and z both belong to $\mathscr{L}^+_{\infty,L_c}(S, \mathscr{S}_0, P)$.

Part (1) is true for all j, since I assume free disposability. That (2) is true follows from assumptions 3.2 and 3.6.

Lemma 8.6 says that $\mathscr{E}_0$ satisfies the boundedness assumption of [5]. Lemmas 8.13 and 8.15 imply that the "adequacy" and "exclusion" assumptions of [5] are satisfied.

The assumptions I have mentioned include all the major assumptions of Theorems 1 and 3 of [5]. It is easy to check that all the minor assumptions are satisfied. These theorems imply that $\mathscr{E}_0$ has an equilibrium. Q.E.D.

10. Proof of Theorem 4.2

I modify the economy $\mathscr{E}$ by replacing the I consumers by a single consumer with utility function $u: R_+^{L_c} \times S \to (-\infty, \infty)$, defined by $u(x, s) = \max\{\sum_{i=1}^I \lambda_i^{-1} u_i(x_i, s) \mid x_i \in R_+^{L_c}$, for all i, and $\sum_{i=1}^I x_i = x\}$. The endowment of the consumer is $\sum_{i=1}^I \omega_i$. Firms remain the same as in $\mathscr{E}$. Call the new economy $\hat{\mathscr{E}}$. It is easy to verify that $\hat{\mathscr{E}}$ satisfies all of assumptions 3.1–3.15. Therefore by Theorem 4.1, $\hat{\mathscr{E}}$ has a stationary equilibrium, $(x, (y_{j0}, y_{j1})_{j=1}^J, p)$. I assume that p is so normalized that the marginal utility of money of the single consumer is one. By applying Aumann's measurable choice theorem [4] to $H = \{(s, x_1, \ldots, x_I) \in S \times (R_+^{L_c} \times \cdots \times R_+^{L_c}) \mid u(x(s), s) =$

$\sum_{i=1}^{I} \lambda_i^{-1} u_i(x_i, s)\}$, one obtains $x_i \in \mathscr{L}^{+}_{\infty, L_c}(S, \mathscr{S}_0, P)$, for $i = 1, \dots, I$, such that $u(x(s), s) = \sum_{i=1}^{I} \lambda_i^{-1} u_i(x_i(s), s)$ almost surely. It follows easily that $((x_i)_{i=1}^{I}, (y_{j0}, y_{j1})_{j=1}^{J}, p)$ is a stationary equilibrium for $\mathscr{E}$ with transfer payments, which satisfies the conditions of the theorem. Q.E.D.

11. Proof of Theorem 4.3

The proof is nearly standard, except that I use the permanent income hypothesis.

Let $((x_i), (y_{j0}, y_{j1}), p)$ be a stationary equilibrium with transfer payments and let λ_i be the marginal utility of money of consumer i in this equilibrium, for $i = 1, \dots, I$. Let $((\bar{x}_i), (\bar{y}_j))$ be a feasible allocation which is not necessarily stationary. (Recall that $\bar{x}_i = (\bar{x}_i^0, \bar{x}_i^1, \dots)$ and $\bar{y}_j = (\bar{y}_{j1}^{-1}, (\bar{y}_{j1}^0, (\bar{y}_{j0}^0, \bar{y}_{j1}^0), (\bar{y}_{j0}^1, \bar{y}_{j1}^1), \dots)$.) Suppose that $\lim_{N \to \infty} \inf N^{-1} \sum_{n=0}^{N} [u_i(\bar{x}_i^n(s), s) - u_i(x_i(\sigma^n s), s)]$ $P(ds) \geqslant 0$, for all i, with strict inequality for some i, say for $i = 1$. I derive a contradiction.

Let $\varepsilon > 0$ be such that $\lim_{N \to \infty} \inf N^{-1} \sum_{n=0}^{N} \int [u_1(\bar{x}_1^n(s), s) - u_1(x_1(\sigma^n s), s)]\, P(ds) \geqslant 3\varepsilon\lambda_1$. Choose $\bar{N}$ so large that if $N \geqslant \bar{N}$, then

$$N^{-1} \sum_{n=0}^{N} \int [u_1(\bar{x}_1^n(s), s) - u_1(x_1(\sigma^n s), s)]\, P(ds) \geqslant 2\varepsilon\lambda_1,$$

and

$$N^{-1} \sum_{n=0}^{N} \int [u_i(\bar{x}_i^n(s), s) - u_i(x_i(\sigma^n s), s)]\, P(ds) \geqslant -I^{-1}\varepsilon\lambda_i, \qquad \text{for } i > 1.$$

Next observe that $u_i(\bar{x}_i^n(s), s) \leqslant u_i(x_i(\sigma^n s), s) + \lambda_i\, p(\sigma^n s) \cdot (\bar{x}_i^n(s) - x_i(\sigma^n s))$, for all i. By combining these inequalities with the previous ones and adding the resulting inequalities, I obtain

$$\sum_{n=0}^{N} \int p(\sigma^n s) \cdot \left[\sum_{i=1}^{I} (\bar{x}_i^n(s) - x_i(\sigma^n s)\right] P(ds) \geqslant \varepsilon N, \qquad \text{if } N \geqslant \bar{N}. \quad (11.1)$$

Since $((\bar{x}_i), (\bar{y}_j))$ and $((x_i), (y_{j0}, y_{j1}))$ are feasible allocations, I obtain $\sum_{n=1}^{I} \bar{x}_i^n = \sum_{i=1}^{I} \sigma^n \omega_i + \sum_{j=1}^{J} (\bar{y}_{j0}^n + \bar{y}_{j1}^{n-1})$, for all n, and $\sum_{i=1}^{I} x_i = \sum_{i=1}^{I} \omega_i + \sum_{j=1}^{J} (y_{j0} + \sigma^{-1} y_{j1})$. By substituting these expressions into (11.1) and rearranging, I obtain

$$\begin{aligned} \sum_{j=1}^{J} \Big\{ & p \cdot (\bar{y}_{j1}^{-1} - \sigma^{-1} y_{j1}) + p \cdot (\sigma^{-N} \bar{y}_{j0}^{N} - y_{j0}) \\ & + \sum_{n=0}^{N-1} [p \cdot (\sigma^{-n} \bar{y}_{j0}^{n} + \sigma^{-(n+1)} \bar{y}_{j1}^{n}) - p \cdot (y_{j0} + \sigma^{-1} y_{j1})] \Big\} \geqslant \varepsilon N, \text{ if } N \geqslant \bar{N}. \end{aligned} \quad (11.2)$$

Since $(y_{j0}, y_{j1}) \in \eta_j(p)$, it follows that $p \cdot (\sigma^{-n}\bar{y}_{j0}^n + \sigma^{-(n+1)}\bar{y}_{j1}^{n+1}) \leqslant p \cdot (y_{j0} + \sigma^{-1}y_{j1})$, for all j and n. Therefore, (11.2) implies that

$$\sum_{j=1}^{J} [p \cdot (\bar{y}_{j1}^{-1} - \sigma^{-1}y_{j1}) + p \cdot (\sigma^{-N}\bar{y}_{j0}^N - y_{j0})] \geqslant \varepsilon N, \quad \text{for all } N \geqslant \bar{N}. \tag{11.3}$$

Since $(\sigma^{-N}y_{j0}^N, \sigma^{-N}y_{j1}^N) \in Y_j$, it follows that $y_{j0}^N \leqslant 0$, for all j. Since $p > 0$, $p \cdot \sigma^{-N}y_{j0}^N \leqslant 0$. Hence, (11.3) implies that $\sum_{j=1}^{J} p \cdot (y_{j1}^{-1} - \sigma^{-1}y_{j1} - y_{j0}) \geqslant \varepsilon N$, for all $N \geqslant \bar{N}$. That is, $\sum_{j=1}^{J} p \cdot (y_{j1}^{-1} - \sigma^{-1}y_{j1} - y_{j0}) = \infty$, which is impossible. This contradiction proves the theorem. Q.E.D.

12. PROOF OF THEOREM 4.4

I call an equilibrium for $\mathscr{E}_0$ a modified equilibrium if the prices belong to $ba_L^+(S, \mathscr{S}_0, P) \times ba_{L_p}^+(S, \mathscr{S}_1, P)$ rather than to $\mathscr{L}_{1,L}^+(S, \mathscr{S}_0, P) \times \mathscr{L}_{1,L_p}^+(S, \mathscr{S}_1, P)$. $(ba_L(S, \mathscr{S}_0, P) \times ba_{L_p}(S, \mathscr{S}_1, P)$ is the set of linear functionals on $\mathscr{L}_{\infty,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}(S, \mathscr{S}_1, P)$ which are continuous with respect to the supremum norm. See Dunford and Schwartz [12, p. 296, Theorem 16].)

I first apply a standard separation argument to obtain prices for the economy $\mathscr{E}_0$. Strong monotonicity implies that I obtain a modified equilibrium for $\mathscr{E}_0$ with transfer payments. By a theorem in [5], I may assume that the prices belong to $\mathscr{L}_1$, so that I can drop the qualifier "modified." The argument of Section 7 then implies that I have a stationary equilibrium of $\mathscr{E}$.

Let $((\bar{x}_i), (\bar{y}_{j0}, \bar{y}_{j1}))$ be a Pareto optimal stationary allocation for $\mathscr{E}$. Let $\bar{y}_{01} = -\sum_{j=1}^{J} \bar{y}_{j1}$ and let $\bar{y}_{00} = -\sigma^{-1}\bar{y}_{01}$. It is easy to see that $((\bar{x}_i, 0), (\bar{y}_{j0}, \bar{y}_{j1})_{j=0}^{J})$ is a Pareto optimal allocation for $\mathscr{E}_0$.

Let $A = \{\sum_{i=1}^{I} (x_i, 0) \mid U_i^0(x_i, 0) > U_i^0(\bar{x}_i, 0)$, for all $i\}$, where U_i^0 is the utility function of consumer i in the economy $\mathscr{E}_0$. Let $B = \sum_{i=1}^{I} (\omega_i, 0) + \sum_{j=1}^{J} Y_j$. The usual argument says that A and B are convex and do not intersect. B contains $\sum_{i=1}^{I} (\omega_i, 0) - (\mathscr{L}_{\infty,L_p}^+(S, \mathscr{S}_0, P) \times \mathscr{L}_{\infty,L_p}^+(S, \mathscr{S}_1, P))$, and so has non-empty interior with respect to the supremum norm, $|\cdot|$. By the separation theorem (Dunford and Schwarz [12, p. 417, Theorem 8]) there exists a non-zero $(p_0, p_1) \in ba_L^+(S, \mathscr{S}_0, P) \times ba_{L_p}^+(S, \mathscr{S}_1, P)$ such that $(p_0, p_1) \cdot (a_0, a_1) \geqslant (p_0, p_1) \cdot (b_0, b_1)$, for all $(a_0, a_1) \in A$ and $(b_0, b_1) \in B$. Clearly, $(p_0, p_1) > 0$.

Since I have assumed that utility functions are monotone, standard arguments prove (12.1) and (12.2) below. (See Debreu [9] or [10].)

$$(\bar{y}_{j0}, \bar{y}_{j1}) \in \eta_j^0(p_0, p_1), \qquad \text{for} \quad j = 0, 1, \ldots, J, \tag{12.1}$$

where η_j^0 is the supply correspondence for firm j in the economy $\mathscr{E}_0$.

$$\text{For all } i,\ p_0 \cdot x_i \geqslant p_0 \cdot \bar{x}_i, \qquad \text{whenever} \quad U_i^0(x_i, 0) > U_i^0(\bar{x}_i, 0). \tag{12.2}$$

I now prove that

$$\text{for all } i,\ p_0 \cdot x_i > p_0 \cdot \bar{x}_i, \qquad \text{whenever} \quad U_i^0(x_i, 0) > U_i^0(\bar{x}_i, 0). \tag{12.3}$$

A standard argument proves the following.

$$\begin{gathered} \text{For all } i, \text{ if } p_0 \cdot \bar{x}_i > 0, \text{ then} \\ p_0 \cdot x_i > p_0 \cdot \bar{x}_i, \qquad \text{whenever} \quad U_i^0(x_i, 0) > U_i^0(\bar{x}_i, 0). \end{gathered} \tag{12.4}$$

(See Debreu [10, "Remark" on p. 591].) I apply (12.4) to prove (12.3).

By Lemma 8.13 there exist $(y_{j0}, y_{j1})_{j=0}^J$ such that $(y_{j0}, y_{j1}) \in Y_j$, for all j, and $\sum_{i=1}^I (\omega_i, 0) + \sum_{j=0}^J (y_{j0}, y_{j1}) \ggg 0$. Since $(p_0, p_1) > 0$, it follows that $(p_0, p_1) \cdot (\sum_{i=1}^I (\omega_i, 0) + \sum_{j=0}^J (y_{j0}, y_{j1})) > 0$. Since the allocation $((\bar{x}_i, 0), (\bar{y}_{j0}, \bar{y}_{j1})_{j=0}^J)$ is feasible and $(\bar{y}_{j0}, \bar{y}_{j1}) \in \eta_j^0(p_0, p_1)$, for all j, I obtain $(p_0, p_1) \cdot (\sum_i (\bar{x}_i, 0)) = (p_0, p_1) \cdot (\sum_i (\omega_i, 0) + \sum_{j=0}^J (\bar{y}_{j0}, \bar{y}_{j1})) \geqslant (p_0, p_1) \cdot (\sum_i (\omega_i, 0) + \sum_{j=0}^J (y_{j0}, y_{j1})) > 0$. Hence for some i, $p_0 \cdot \bar{x}_i > 0$, say for $i = 1$. By (12.4), (12.3) is true for $i = 1$.

Let $(p_0, p_1) = (p_{0c}, p_{1c}) + (p_{0p}, p_{1p})$ be the Yosida–Hewitt decomposition of (p_0, p_1). (This is defined in the subsection of section 8 dealing with the exclusion assumption.) By the Radon–Nikodyn theorem (Halmos [18, p. 128]), (p_{0c}, p_{1c}) may be thought of as belonging to $\mathscr{L}_{1,L}(S, \mathscr{S}_0, P) \times \mathscr{L}_{1,L_p}(S, \mathscr{S}_1, P)$.

Because (12.3) is true for $i = 1$, it follows easily from the strong monotonicity assumption (3.16) that $p_{0c} \gg 0$. This in turn implies that if $\bar{x}_i > 0$, then $p_0 \cdot \bar{x}_i > 0$. Hence by (12.4), (12.3) is true for all i such that $\bar{x}_i > 0$. If $\bar{x}_i = 0$, then $p_0 \cdot x_i > 0$ for any $x_i \in \mathscr{L}_{\infty,L_c}^{+}(S, \mathscr{S}_0, P)$ not equal to $\bar{x}_i$. Hence, (12.3) is trivially true in this case. This proves (12.3).

Equations (12.1) and (12.3) together imply that $((\bar{x}_i, 0), (\bar{y}_{j0}, \bar{y}_{j1})_{j=0}^J, (p_0, p_1))$ is a modified equilibrium for $\mathscr{E}_0$ with transfer payments. Theorem 3 of [5] implies that $((\bar{x}_i, 0), (\bar{y}_{j0}, \bar{y}_{j1})_{j=0}^J, (p_{0c}, p_{1c}))$ is an equilibrium for $\mathscr{E}_0$ with transfer payments. By the argument of Section 7, $((\bar{x}_i), (\bar{y}_{j0}, \bar{y}_{j1})_{j=1}^J, p_{0c})$ is a stationary equilibrium for $\mathscr{E}$ with transfer payments. Q.E.D.

REFERENCES

1. B. Allen, "Generic Existence of Equilibria for Economics with Uncertainty when Prices Convey Information," Working Paper IP-265, Center for Research in Management Science, University of California, Berkeley, 1978.
2. K. Arrow, The role of securities in the optimal allocation of risk-bearing, *Rev. Econ. Studies* 31 (1963–1964), 91–96.

3. K. Arrow, Limited knowledge and economic analysis, *Amer. Econ. Rev.* **64** (1974), 1–10.
4. R. Aumann, Measurable utility and the measurable choice theorem, *in* "La Décision," pp. 15–26, Editions du Centre National de la Recherche Scientifique, Paris, 1969.
5. T. Bewley, Existence of equilibria in economics with infinitely many commodities, *J. Econ. Theory* **4** (1972), 514–540.
6. T. Bewley, The permanent income hypothesis: A theoretical formulation, *J. Econ. Theory* **16** (1977), 252–292.
7. W. Brock and L. Mirman, Optimal economic growth and uncertainty: The no discounting case, *Internat. Econ. Rev.* **14** (1973), 560–573.
8. R. A. Dana, Evaluation of development programs in a stationary stochastic economy with bounded primary resources, *in* "Proceedings, Warsaw Symposium on Mathematical Methods in Economics" (Jerzy Lós, Ed.), North-Holland, Amsterdam, 1973.
9. G. Debreu, Valuation equilibrium and Pareto optimum, *Proc. Nat. Acad. Sci. USA* **40** (1954), 588–592.
10. G. Debreu, "Theory of Value" Wiley, New York, 1959.
11. J. L. Doob, "Stochastic Processes" Wiley, New York, 1963.
12. N. Dunford and J. Schwartz, "Linear Operators, Part I," Interscience, New York, 1957
13. I. V. Evstigneev, Optimal stochastic programs and their stimulating prices, *in* "Mathematical Methods in Economics" (J. Lós and M. W. Lós, Eds.), pp. 219–252, North-Holland, Amsterdam, 1974.
14. M. Friedman, "A Theory of the Consumption Function" Princeton Univ. Press, Princeton, N. J., 1957.
15. C. Futia, "Rational Expectations in Speculative Markets," Working Paper, Bell Telephone Laboratories, Murray Hill, N. J. 07974, 1979.
16. J. M. Grandmont, Temporary general equilibrium theory, *Econometrica* **45** (1977), 535–572.
17. J. Green, The non-existence of informational equilibria, *Rev. Econ. Studies* **44** (1977), 451–463.
18. P. Halmos, "Measure Theory" Van Nostrand, Princeton, N. J., 1950.
19. P. Jeanjean, Optimal development programs under uncertainty: The undiscounted case, *J. Econ. Theory* **7** (1974), 66–92.
20. J. S. Jordan, Expectations equilibrium and informational efficiency for stochastic environments, *J. Econ. Theory* **16** (1977), 354–372.
21. D. Kreps, A note on 'Fulfilled expectations' equilibria *J. Econ. Theory* **14** (1977), 32–43.
22. E. Malinvaud, The allocation of individual risks in large markets, *J. Econ. Theory* **4** (1972), 312–328.
23. J. Muth, Rational expectations and the theory of price movements, *Econometrica* **29** (1961), 315–335.
24. R. Radner, "Existence of Equilibrium Plans, Prices, and Price Expectations in a Sequence of Markets," *Econometrica*, 40 (1972), 289–303.
25. R. Radner, Optimal stationary consumption with stochastic production and resources, *J. Econ. Theory* **6** (1973), 68–90.
26. R. Radner, Market equilibrium under uncertainty: Concepts and problems, *in* "Frontiers of Quantitative Economics, Vol. II" (M. D. Intrilligator and D. A. Kendrick, Eds.), Chap. 2, North-Holland, Amsterdam, 1974.
27. R. Radner, Rational expectations equilibrium: Generic existence and information revealed by prices, *Econometrica* **47** (1979), 655–678.
28. R. Shiller, Rational expectations and the dynamic structure of macroeconomic models. a critical review, *J. Monetary Econ.* **4** (1978), 1–44.

29. B. Stigum, Competitive equilibria with infinitely many commodities, *Metroeconomica* **24** (1972), 221–244.
30. B. Stigum, Competitive equilibria with infinitely many commodities, II, *J. Econ. Theory* **6** (1973), 415–445.
31. K. Yosida and E. Hewitt, Finitely additive measures, *Trans. Amer. Math. Soc.* **72** (1956), 46–66.

JOURNAL OF ECONOMIC THEORY 24, 296–309 (1981)

Notes, Comments, and Letters to the Editor

Optimal Growth with Experimentation

XAVIER FREIXAS

Université des Sciences Sociales de Toulouse,
Place Anatole France, 31070 Toulouse, France

Received September 5, 1979; revised February 13, 1980

One sector optimal growth models have been widely analyzed within a deterministic framework (see [7] and its references). More recently their study has been extended to models embodying a random technology [2, 10–12]. Still, all these models assume a perfect knowledge of the distribution of the random technology. If we assume imperfect information instead, two economically meaningful issues have to be analyzed. First, the planning authority (P.A.) knows he will have better information about the distribution of the randomness in subsequent periods (we will refer to this phenomenon as "learning") and this will generally affect the first period consumption-investment decision. Second, the first period decision may have an informational effect; some decisions may reduce the P.A. (subjective) uncertainty faster than others. This is so when the P.A. subjective uncertainty is instrument depending. A rational P.A. will take this dependence into consideration and will "experiment" by slightly modifying the first period decision so as to obtain a better future information.

Our aim here is to analyze the direction of the shifts introduced by learning and experimenting in an optimal growth path.

In order to give some intuition on the effect of learning and experimentation it is interesting to recall the results obtained by Grossman, Kihlstrom and Mirman (GKM) [6], in a related context.

GKM studied the effect of experimentation for a consumer who learns about the random effect of a drug on his health. The more drug he uses the better he knows the distribution of the randomness, i.e., the more he learns, and therefore experimentation increases the first period drug consumption. In other words, the experimenting consumer chooses to invest on learning (by consuming more of the drug) during the first period. And this result extends easily to a wide range of economic models involving instrument depending randomness: an experimenting decision-maker will choose to invest on learning, so that the first period decision of his optimal growth path (when

0022-0531/81/020296–14$02.00/0

1981 by Academic Press, Inc.

compared with the one of the non learning decision-maker) shifts towards decisions yielding better future information structures.

How would the GKM result apply to optimal growth? Suppose that we learn more about the unknown random production function when we invest more. Then, according to GKM, experimentation will increase the investment chosen during the first period. But there is a difference to be underlined between the GKM model and an optimal growth model. In GKM a faster learning, i.e., a faster resolution of the subjective uncertainty has no effect on the first period decision, while, as we will show, it shifts the first period decision in an optimal growth model.

Consequently two cases have to be distinguished. If learning increases the expected marginal return of investment, then its effect will be to decrease the level of the first period optimal consumption. Conversely, if learning decreases the expected marginal return of investment, it will increase the level of the first period optimal consumption. Only in the first case can we assert that the GKM result will hold.

On the other hand, we prove that the departure of the experimenting growth path from the learning growth path is always in the direction determined by the GKM analysis: it increases the first period's investment. But as learning may decrease the investment, the total effect of experimentation when compared to non-learning is ambigous.

For the sake of completeness, we first describe, in Section 1, the classical model of stochastic growth without learning. Section 2 is devoted to the introduction of learning and experimentation in the model. In Sections 3 and 4 we develop the comparative dynamics of learning and experimenting.

1. Optimal Growth without Learning

We consider a simple, one good, optimal stochastic growth model where the only randomness appears in the production function. The output $\tilde{k}_t$ of the only good in period t is a random variable which depends on the input x_{t-1} and on a random variable ω_t.

We write f the production function

$$\tilde{k}_t = f(x_{t-1}, \omega_t). \tag{1.1}$$

In every period t, the output will be divided between the consumption C_t and the investment x_t which will yield a random output $\tilde{k}_{t+1}$ in the next one.

We assume there is a stationary utility function $U(\cdot)$ defined on $\mathbb{R}^+$ and a discount factor $\delta (0 < \delta < 1)$ so that the value associated to the sequence of positive consumptions $C_0, C_1, \dots, C_T$ will be $\sum_{t=0}^{T} \delta^t U(C_t)$, where T is the planning horizon.

We will make the following classical assumptions concerning the production function and the utility function.

A1. For every ω_t, the function $f(\cdot, \omega_t)$ is positive with

$$\frac{\partial f}{\partial x_{t-1}}(\cdot, \omega_t) > 0, \qquad \frac{\partial^2 f}{\partial x_{t-1}^2}(\cdot, \omega_t) \leqslant 0 \qquad \text{and} \qquad f(0, \omega_t) = 0.$$

A2. We assume U is a twice differentiable function satisfying $U'(0) = +\infty$, $U' > 0$, $U'' < 0$.

We also assume that the distribution of ω_t is stationary. We will write its density $\psi(\omega_t, p)$, where p indicates the information we have about the distribution of ω_t. Namely, p is the P.A.'s subjective probability distribution. (In this section p is only introduced in order to have notations consistent with the problems introduced below.)

Under the above assumptions, the optimal stochastical growth path is defined as the (random) sequence $C_0^*, C_1^*(\omega_1), \ldots, C_T^*(\omega_1, \omega_2, \ldots, \omega_T)$ solution of

$$\begin{aligned} &\operatorname{Max} E \left\{ \sum_{t=0}^{T} \delta^t U(C_t) \right\}, && \\ &C_0 + x_0 = S, && \\ &C_t + x_t = f(x_{t-1}, \omega_t), && t = 1, 2, \ldots, T, \\ &C_t \geqslant 0, && t = 0, \ldots, T, \\ &x_t \geqslant 0, && t = 0, \ldots, T. \end{aligned} \tag{1.2}$$

It is easy to show that for every p and t there exists a continuous concave value function $V^t(S, p)$ equal to the solution value of (1.2) for $t = T$. The optimality principle implies

$$V^T(S, p) = \operatorname*{Max}_{0 \leqslant C \leqslant S} U(C) + \delta E[V^{T-1}(f(S - C, \omega), p)], \tag{1.3}$$

where ω stands for the stationary random variable. The solution of this problem depends on t through the quantity of good inherited from the past, and through the horizon. The optimal policies at period $T-n$ are. $C_{T-n}^* = g^n(S, p)$, where g^n is an increasing function of S (see [11]). Therefore, in order to implement the optimal consumption for period t, we need not know $(\omega_1, \omega_2, \ldots, \omega_T)$, but just S.

Using the envelope theorem it is possible to show that under the above assumptions V^T is differentiable with respect to S and that we have

$$V_s^T(S, p) = U'(g^T(S, p)). \tag{1.4}$$

Therefore we obtain the first order condition for the maximization problem (1.3)

$$U'(C) - \delta E\{U'[g^{T-1}(f(S-C,\omega), p]f'(S-C,\omega)\} = 0 \qquad (1.5)$$

and the solution $C^* = g^T(S, p)$ of (1.5) is the optimal consumption for the current period.

2. LEARNING AND EXPERIMENTATION

For the sake of simplicity we begin by discussing an example of learning. Assume the unknown random production function is

$$f(x, A, \omega) = Ax + \varepsilon(\omega), \qquad (2.1)$$

where A is a certain value. The P.A. does not know the actual value of A but he knows A belongs to a set $\{A_1, A_2, \ldots, A_n\}$ and he has a subjective probability distribution for A. That is, he has estimates for the probability of the event $\{A = A_l\}$ we will denote p_l^0.

After producing some output k, information is available since we have a (one item) sample of the random variable $\tilde{k} = f(x, A, \omega)$.

A non-learning P.A. will disregard the new information and keep the same estimates $(p_1^0, p_2^0, \ldots, p_n^0)$ of the probabilities of the different outcomes; he will solve the problem as in the preceding section, since his estimates are stationary. Note that in this context this is naive (and clearly nonrational) behavior.

On the other hand, a learning P.A. will instead make use of the information to revise his estimates $p^0 = (p_1^0, p_2^0, \ldots, p_n^0)$ by means of Bayes' rule.

If we assume, for simplicity sake[1] (at the expense of having a negative

---

[1] For the distribution of $\varepsilon(\omega)$ yielding only positive productions, this random variable is bounded from below and some values of k are incompatible with some values A_l. The updating rule becomes then more complex and has different analytical expressions depending on the value k of $\tilde{k}$ obtained. For instance, if we choose an $\varepsilon(\omega)$ uniformly distributed between zero and one, the updating rule for $A = A_1$ will be

$$p_1^1 = 0 \qquad \text{if} \quad k \notin [A_1x, A_1x + 1],$$
$$p_1^1 = \frac{p_1^0}{\sum_{l \in I(k)} p_l^0} \qquad \text{if} \quad k \in [A_1x, A_1x + 1],$$

with $I(k) = \{i \mid k \in [A_ix, A_ix + 1]\}$.

Showing that larger levels of input yield a better information becomes then more involved without adding to the understanding of the underlying economics.

production with non-zero probability) that $\varepsilon(\omega)$ is normal with a zero mean and variance 1, the updating rule will be given by

$$p_i^1 = P(A = A_i/k, x) = \frac{\chi(k/xA_i)\, p_i^0}{\sum_{j=1}^{n} p_j^0 \chi(k/xA_j)}, \tag{2.2}$$

where $\chi(\ |\mu)$ stands for the density of a normal variable with mean μ and variance 1.

The posterior probabilities will depend both on ω and on x, and x is chosen by the P.A. A learning P.A. will make use of Bayes' rule to update his estimates of the event $A = A_i$. Therefore, he knows in period t that he will face a different probability distribution for A in period $t+1$ even if he still does not know which one. Consequently, he will choose his first period decision so as to take into account that a better information will be available in the future (Note that the first period optimal decision is taken before any observation of k by the P.A. takes place.) However, a learning P.A. is oblivious to the fact that his decision generates information.

On the other hand, an experimenting P.A. will notice the effect of the input level x on the information he will have concerning A. In this example an increase in the input level x generates a better information structure. This becomes clear when we consider the auxiliary random variable $\tilde{r} = \tilde{k}/x$ which is normally distributed with mean A and variance $1/x^2$. A larger x is therefore "more informative" since the observed values of r will have the same unknown mean A and a smaller variance. Consequently, here, an experimenting P.A. will tend to choose higher levels of input than a learning one.

In a more general context, production $\tilde{k}$ will depend on the input level x, the random variable $\varepsilon(\omega)$ and the unknown parameter A. We will denote Bayes' rule in vector notations.

$$p^1 = \varphi(p^0, k, x) \qquad \text{with} \qquad p^1 = (p', p'_2, \ldots, p_n^1). \tag{2.3}$$

The behavior of a learning P.A. discussed above can be summarized in the expression he uses to update his estimates

$$p^t = \tilde{\varphi}(p^{t-1}, k_t) \qquad \text{for} \quad t = 1, \ldots, T, \tag{2.4}$$

while an experimenting P.A. will consider

$$p^t = \varphi(p^{t-1}, k_t, x_t). \tag{2.5}$$

Of course, for consistency, we require

$$\tilde{\varphi}(p^{t-1}, k_t) = \varphi(p^{t-1}, k_t, \hat{x}_{t-1}), \qquad (2.6$$

where $\hat{x}_{t-1}$ is the decision taken in period $t-1$ by a learning P.A.

Explicitly, formulas (2.4) and (2.5) are given, respectively, by

$$p_i^t = \frac{\theta(k \mid A_i)\, p_i^{t-1}}{\sum_{j=1}^{n} p_j^{t-1}\theta(k \mid A_j)} \tag{2.7}$$

for a learning P.A. and

$$p_i^t = \frac{\theta(k \mid x, A_i)\, p_i^{t-1}}{\sum_{j=1}^{n} p_j^{t-1}\theta(k \mid x, A_j)} \tag{2.8}$$

for an experimenting P.A., where $\theta(k \mid A_i)$ is the generalized probability density function (g.p.d.f.) of the random variable $\tilde{k}$ for given A_i, and equivalently, $\theta(k \mid x, A_i)$ is the g.p.d.f. for a given x and A_i.

Note that before observing k_t, p^t is a random variable.

We will assume that larger amounts of the input x perform an experiment $(\tilde{k}, x)$ which is sufficient for the experiment yielded by smaller amounts of input. That is to say, taking observations of the random variable $\tilde{k}$ is more informative for larger values of x. Formally

A3. If $x' > x$, then there is a positive function $h(k, k')$ such that

(i) $\theta(k \mid x, A_i) = \int_{-\infty}^{+\infty} h(k, k')\, \theta(k' \mid x', A_i)\, dk'$,

(ii) $\int_{-\infty}^{+\infty} h(k, k')\, dk = 1$.

Under assumption A3, we will say that larger amounts of input are "more informative."

Assumption A3 can be shown to hold for the above example (see [3, 9]).

In an optimal growth context a learning or experimenting P.A. will face a different future random probability distribution in every period.

The optimal growth path will be the solution of

$$\operatorname{Max} U(C_0) + \delta E[U(C_1) \mid p^0] + \sum_{t=2}^{T} \delta^t E[U(C_t) \mid p^{t-1}]$$

$$\begin{array}{ll} C_0 + x_0 = S, & \\ C_t + x_t = \tilde{k}_t, & t = 1, \ldots, T \\ \tilde{k}_t = f(x_{t-1}, A, \varepsilon(\omega_t)), & t = 1, \ldots, T, \\ x_t \geqslant 0;\ C_t \geqslant 0, & t = 0, \ldots, T, \end{array} \tag{2.9}$$

where p^{t-1} is given by formula (2.7) for a learning P.A. and by formula (2.8) for an experimenting P.A.

We note $V^T(S, p^0)$ the maximum value of the objective function for this problem. If V^T exists the optimality principle can then be written for a learning P.A. as

$$V^T(S, p^0) = \operatorname*{Max}_{0 \leqslant C \leqslant S} U(C) + \delta E[V^{T-1}(f(S-C, A, \varepsilon(\omega)), \bar{\varphi}(p^0, \tilde{k}_1)] \quad (2.10)$$

with $\tilde{k}_1 = f(S - C, A, \varepsilon(\omega)$ and it will be written

$$V^T(S, p^0) = \operatorname*{Max}_{0 \leqslant C \leqslant S} U(C) + \delta E[V^{T-1}(f(S-C, A, \varepsilon(\omega)), \varphi(p^0, \tilde{k}_1, S - C)] \quad (2.11)$$

for an experimenting P.A.

3. THE COMPARATIVE DYNAMICS OF LEARNING

In this section we restrict our attention to learning and compare the decisions taken by a learning and a non-learning P.A. so as to identify the effect of a better future information structure on the first period consumption.

Since the introduction of learning in the optimal growth model modifies the value function, our first step will be to prove the following proposition.

PROPOSITION 1. *Under assumptions* A1, A2, $V^n(S, p)$ *exists and is continuously differentiable for every* n, $0 \leqslant n \leqslant T$.

Proof. The proof is accomplished by using an induction argument.

First, for $n = 0$,

$$V^0(S, p) = \operatorname*{Max}_{0 \leqslant C \leqslant S} U(C).$$

as U is increasing, the optimal policy, $g^0(S, p) = S$, is differentiable. In addition, $g^0_S(S, p) > 0$. Consequently

$$V^0(S, p) = U(g^0(S, p))$$

also exists and is differentiable.

Assume now that g^{n-1} and V^{n-1} exist, are continuously differentiable and $g^{n-1}_S > 0$.

Note $\hat{C} = g^n(S, p)$. It is easy to prove that $\hat{C}$ exists using the continuity of V^{n-1}. Furthermore, as $V'(0) = +\infty$ we will not have any corner solutions, since this would imply an infinite marginal utility either for period n or for the following one $(n - 1)$.

Consequently,

$$0 < \hat{C} < S.$$

Define the function W by

$$W(C, S, p) = U'(C) - \delta E[U'(g^{n-1}(\bar{k}, \bar{\varphi}(p, \bar{k}) f'(S - C, A, \varepsilon(\omega)].$$

We have

$$W(\hat{\hat{C}}, S, p) = 0.$$

Since U' and g^{n-1} are continuously differentiable, W is differentiable. In addition, the partial derivative with respect to C is[2]

$$W_C(C, S, p) = U''(C) + \delta E\{[U''(g^{n-1}(\bar{k}, \bar{\varphi}(p, \bar{k}))\, g_S^{n-1}(\bar{k}, \bar{\varphi}(p, k) f'(S - C, \omega)] \\ + U'(g^{n-1}(\bar{k}, \bar{\varphi}(p, \bar{k}))\, f''(S - C, \omega)\}$$

with $U'' < 0$ and $g_s^{n-1} > 0$, so that $W_C(C, S, p)$ is negative.

We can therefore apply the implicit function theorem (see [4, 10.2]) and show $g^n(S, p)$ is a continuously differentiable function of S and p.

In addition, note that $g_s^n(S, p) = -W_s/W_c$ is positive since $W_s > 0$ and $W_c < 0$.

Finally we can write

$$V^n(S, p) = U(g^n(S, p)) + \delta E\{V^{n-1}[f(S - g^n(S, p), A, \varepsilon(\omega); \bar{\varphi}(p, \bar{k})]\}.$$

V^{n-1} is by assumption continuously differentiable and so are g^n, f and the operator E. As $\bar{\varphi}$ has continuous partial derivatives with respect to p for every k, V^n is differentiable.

Proposition 1 authorizes us to set the first order conditions of problem (2.10)

$$U'(C) - \delta E\{V_S^{T-1}(\bar{k}, \bar{\varphi}(p, \bar{k}))\, f'(S - C, A, \varepsilon(\omega)\} = 0 \tag{3.1}$$

and using (1.4) we can rewrite it as

$$U'(C) - \delta E\{U'(g^{T-1}(\bar{k}, \bar{\varphi}(p, \bar{k})\, f'(S - C, A, \varepsilon(\omega)\} = 0. \tag{3.2}$$

In order to compare the solution $\hat{C}$ of the non-learning problem (2.1) with the solution $\hat{\hat{C}}$ of (3.2), we compute

$$D_L(\hat{\hat{C}}) = U'(\hat{\hat{C}}) - \delta E\{U'(g^{T-1}(\bar{k}, p)\, f'(S - \hat{C}, A, \varepsilon(\omega)\}. \tag{3.3}$$

As the problem (1.2) is concave $D_L(\cdot)$ is a decreasing function. Thus we can infer from the sign of $D_L(\hat{\hat{C}})$ whether $\hat{\hat{C}}$ is superior or inferior to $\hat{C}$: if $D_L(\hat{\hat{C}})$ is negative, $\hat{C}$ is superior to $\hat{\hat{C}}$ and reciprocally.

---

[2] Note that learning implies $d\varphi(p, k)/dc = d\varphi(p, k)/ds = 0$.

Using (3.2) we can obtain an equivalent expression for D_L:

$$D_L(\dot{\hat{C}}) = \delta\{E[U'(g^{T-1}(\tilde{k}, \bar{\varphi}(p, \tilde{k})\, f'(S - \dot{\hat{C}}, A, \varepsilon(\omega))] \\ - E[U'(g^{T-1}(\tilde{k}, p)\, f'(S - \dot{\hat{C}}, A, \varepsilon(\omega))]\}. \qquad (3.4)$$

The RHS of (3.4) is δ times the expected difference between the marginal return of investment with learning and the marginal return of investment, without learning. Consequently, if learning increases the marginal return of investment, $D_L(\dot{\hat{C}})$ will be positive, and $\dot{\hat{C}}$ will be inferior to $\hat{C}$. But if learning decreases the marginal return of investment, the opposite will be true.

The rest of this section is devoted to showing that in a stochastic growth model, the introduction of learning may either increase or decrease the first period consumption.

We distinguish two different kinds of situations. For a given S and for a given learning structure, we will say that the couple (u, f) of utility and production functions is LI (respectively LD) if learning increases (resp. decreases) the marginal return of investment, i.e.,

$$E[U'(g^{T-1}(\tilde{k}, \bar{\varphi}(p, \tilde{k})\, f'(S - \dot{\hat{C}}, A, \varepsilon(\omega)) \\ - U'(g^{T-1}(\tilde{k}, p)\, f'(S - \dot{\hat{C}}, A, \varepsilon(\omega))] \geqslant 0 \qquad (\text{resp.} \leqslant 0). \qquad (3.5)$$

We can therefore summarize the above discussion in the following proposition.

PROPOSITION 2. *If* (U, f) *are* LI (*respectively* LD), *then* $\dot{\hat{C}} \leqslant \hat{C}$ (*resp.* $\dot{\hat{C}} \geqslant \hat{C}$).

We cannot, a priori, give an explicit definition of the class of LI couples of functions. Yet we can establish an interesting sufficiency criterion.

PROPOSITION 3. *If* V^{T-1} *is concave* (*respectively convex*) *with respect to its second argument* p, *then* (U, f) *is* LD (*resp.* LI).

Proof. Note first that the experiment $\tilde{k}$ performed by a learning P.A. is sufficient for the "experiment" performed by a non-learning P.A. since we may think of a non-learning P.A. as if he were observing an experiment k_0 which gives no information about A. In this case we have $\theta(k_0 \mid A_i) = \lambda(k_0)$ for every i, and we obtain

(i) $\theta(k_0 \mid A_i) = \int_{-\infty}^{+\infty} \lambda(k_0)\, \theta(k \mid A_i)\, dk$,

(ii) $\int_{-\infty}^{+\infty} \lambda(k_0)\, dk_0 = 1$.

We can therefore apply Blackwell's theorem to the uncertainty function V_S^{T-1} (See [1, 3, 8, 9]). We then use (1.4) to obtain the formula opposite to (3.5).

Finally, we may wonder whether the assumptions A1, A2 on the utility and production functions may imply their being LI or LD. This possibility has not been warded off, since we have not given an explicit definition of the sets of LI or LD functions.

Applying Proposition 3 to the following example shows how functions U and f satisfying A1 and A2 can be either LI or LD.

EXAMPLE. We consider a three period problem ($T = 2$).

We take $f(x, A, \varepsilon(\omega)) = [A + \varepsilon(\omega)]x$.

Here two experiments performed with input levels x and x' yielding the same information, since $f(x, A, \varepsilon(\omega))/x$ and $f(x', A', \varepsilon(\omega)/x'$ are identical.

At the end of period zero, we receive some information and take p_i^1 as the probability distribution for period 1. We write $\tilde{p}_i^1$ for the random variable whose realizations are the probabilities p_i^1.

We assume $U(x) = x^{1-\gamma}/1 - \gamma$, $\gamma > 0$, $\gamma \neq 1$.

U and f clearly satisfy assumptions A1 and A2.

Making use of Proposition 3, we will show that (U, f) are LI or LD depending on the value of γ. In order to do so we compute $V_S^1(k^1, p^1)$, where k^1 is the quantity of good obtained in period 1.

For a given k^1 and p^1 the problem faced by the P.A. is

$$\operatorname*{Mac}_{0 \leqslant C \leqslant k^1} U(C) + \delta \sum_{i=1}^{n} p_i^1 U(A_i(k^1 - C))$$

and the first order conditions for this problem are

$$C^{-\gamma} - \delta \sum_{i=1}^{n} p_i^1 (A_i(k^1 - C))^{-\gamma} A_i = 0.$$

The solution $\hat{C}$ can easily by obtained

$$\hat{C} = \frac{k^1 \alpha^{-1/\gamma}}{1 + \alpha^{-1/\gamma}}, \qquad \text{with } \alpha = \delta \sum_{i=1}^{n} p_i^1 A_i^{1-\gamma}.$$

Now using (1.4) we obtain

$$V_S^1(k^1, p^1) = \frac{(k^1)^{-\gamma} \alpha}{[1 + \alpha^{-1/\gamma}]^{-\gamma}}.$$

To prove that V_S^1 is a concave or convex function of p^1 it is sufficient to show that it is a concave or convex function of α, since α is a linear function of p_i^1.

We therefore compute the second derivative of V_S^1 as a function of α and obtain

$$\frac{\partial^2 V_S^1}{\partial \alpha^2} = (k_1)^{-\gamma} \frac{(1-\gamma)}{\gamma} [1 + \alpha^{-1/\gamma}]^{\gamma-2} \alpha^{-1/\gamma-1}.$$

Consequently if $\gamma < 1$, V_S^1 is convex and (u, f) are LI. On the other hand if $\gamma > 1$, (U, f) are LD. Therefore if the relative risk aversion measure is inferior to one, the prospect of having more information will decrease current consumption and reciprocally.

It is interesting to compare this result with the effect of an increase in risk on the first period sharing between consumption and investment. In a stochastic growth context two kinds of responses to a more risky environment are possible. First, the optimal policy may involve a larger current consumption, since this yields some utility with certainty. Second it may entail a larger investment in order to prevent a low production level in the future. The choice between these two policies will depend on the utility and production functions (see also [5]).

A quite similar phenomenon occurs when we compare different future information structures. A better future information structure would be the equivalent of a less risky environment, and the response to the change in future information may imply either an increase or a decrease of current consumption.

4. THE COMPARATIVE DYNAMICS OF EXPERIMENTATION

We will show here that the GKM result applies when we compare experimentation and learning while it generally does not hold when we compare experimentation and non-learning. As under assumption A3 higher levels of input are more informative, we first want to establish that an experimenting P.A. will invest more than a learning one.

If the experimenting problem was differentiable, it whould be sufficient to show that the derivative of the objective function (2.11) with respect to consumption, is negative at the point that a learning P.A. would choose. But without additional assumptions the experimenting problem may be non-differentiable[3] if the objective function is non-concave.

Consequently, in order to avoid innecessary assumptions we will dispense with differentiation and show an equivalent property: evaluated at the point a learning P.A. would choose, the objective function of an experimenting P.A.

---

[3] A sufficient condition for differentiability would be to assume that the objective function is a Morse function.

equals the objective function of a learning P.A. plus a decreasing function of the consumption level C.

In other words, we want to show that the difference between the objective functions (2.11) and (2.10).

$$\begin{aligned} H(C) = E\{ & V^{T-1}(\bar{k}, \varphi(p^0, f(S-C, A, \varepsilon(\omega)), S-C) \\ & - V^{T-1}(\bar{k}, \bar{\varphi}(p^0, f(S-C, A, \varepsilon(\omega))\} \end{aligned}$$

is locally non-increasing at the point $\hat{C}$.

Proposition 5. *Under* A1, A2, A3 *any solution* $\bar{C}$ *of the experimenting problem is inferior or equal to the solution* $\hat{C}$ *of the learning problem.*

Proof. We will show that the difference between the objective functions (2.11) and (2.10).

$$\begin{aligned} H(C) = E\{ & V^{T-1}(\bar{k}, \varphi(p^0, f(S-C, A, \varepsilon(\omega)), S-C) \\ & - V^{T-1}(\bar{k}, \bar{\varphi}(p^0, f(S-C, A, \varepsilon(\omega))\} \end{aligned}$$

is a non-increasing function of C at the point $\hat{C}$.

Note first that $H(\hat{C}) = 0$ because of (2.6). Now, under assumption 3, smaller levels of C (i.e., larger investments) are more informative.

As we have

$$\begin{aligned} V^{T-1}(k, p) = \operatorname{Max} E \Bigg\{ & \sum_{t=0}^{T-1} \delta^t V(C_t) \mid C_t + x_t = f(x_{t-1}, A, \varepsilon(\omega)) \\ & C_0 + x_0 = k, C_t \geqslant 0, x_t \geqslant 0 \Bigg\} \end{aligned}$$

we can apply Blakwell's theorem (see, for example, Theorem 6.3 in [9]) to prove that

$$E\{V^{T-1}(\bar{k}, \varphi(p, f(S+\Delta-\hat{C}, A, \varepsilon(\omega)), S+\Delta-\hat{C}) - V^{T-1}(\bar{k}, \varphi(p, \bar{k}, S-\hat{C})\}$$

has the same sign than Δ. As $\bar{\varphi}(p, \bar{k}) = \varphi(p, S-\hat{C})$, this implies $H(\hat{C}-\Delta)$ has the same sign as Δ, so that H is locally a nonincreasing function.

The comparison of the solution $\bar{C}$ of the experimenting with the solution $\hat{C}$ of the non-learning one becomes very easy when combining Propositions 2 and 5:

Proposition 6. *If* (U, f) *are* LI *for the equivalent learning problem,* $\hat{C}$ *is inferior to* $\bar{C}$.

The proof is obvious.

On the other hand, if (U, f) are LD, the solution of the experimenting problem is ambiguous.

In the context, the distinction between the learning effect and the experimenting effect brings in a better understanding of the underlying economic phenomena. Considered separately, the first effect would induce an increase in current consumption and the second a decrease. The joint effect will clearly depend on both. If the experimenting effect dominates the learning effect, $\bar{C}$ will be inferior to $\hat{C}$. Otherwise, if the learning effect dominates the experimenting effect, $\bar{C}$ will be superior to $\hat{C}$.

To conclude we summarize our results emphasizing the difference with those of GKM.

Proposition 5 is very close to GKM's conclusions. It shows that the shift in the optimal growth path due to the introduction of experimentation in a learning process will increase investment. As we have assumed that a larger investment yields a better information, this means that the optimal growth path will be displaced towards more informative decisions in the first period.

But if we compare experimentation with non-learning the conclusion depends on the effect of learning. For the class LI of utility and production functions an improvement in future information will bring about a larger investment, and this will yield a better information. In this case the two effects add and the consumption decreases. On the other hand, for the class LD of utility and production functions, the prospect of a better future information implies a decrease in the investment. Therefore investment will increase only if the experimenting effect goes beyond the learning effect, and reciprocally.

Finally note that the comparison between non-learning and experimenting is the relevant one, (and this is what is compared in GKM). Indeed a learning and non-experimenting decision maker would not be rational i experimenting yields an increase in the total discounted utility obtained fo the period.

REFERENCES

1. D. Blackwell, Equivalent comparison of experiments, *Ann. Math. Statist.* (195 265–272.
2. W. Brock and L. Mirman, Optimal economic growth and uncertainty: The discoun case, *J. Econ. Theory* 4(1972), 479–513.
3. M. DeGroot, "Optimal Statistical Decisions," McGraw–Hill, New York, 1970.
4. J. Dieudonne, "Foundations of Modern Analysis," Academic Press, New York, 196
5. X. Freixas and R. Kihlstrom, A note on information in the market for doct unpublished paper, University of Illinois at Urbana–Champaign, November 1978.
6. S. Grossman, R. Kihlstrom and L. Mirman, A Bayesian approach to the produc of information and learning by doing, *Rev. Econ. Studies* **44** (1977), 533–547.

7. M. D. INTRILLIGATOR, "Mathematical Optimization and Economic Theory," Prentice–Hall, Englewood Cliffs, N. J., 1971.
8. R. KIHLSTROM, A Bayesian exposition of Blackwell's theorem on the comparison of experiments, unpublished paper, State University of New York at Stony Brook, March 1974.
9. J. MARSCHAK AND K. MIYASAWA, Economic comparibility of information systems, *Internat. Econ. Review* (June 1968).
10. L. MIRMAN, The steady state behavior of a class of one sector growth models with uncertain technology, *J. Econ. Theory* 5 (1973), 219–242.
11. L. MIRMAN, One sector economic growth and uncertainty: A survey, unpublished paper, University of Illinois at Urbana–Champaign, 1976.
12. L. MIRMAN AND I. ZILCHA, On optimal growth under uncertainty, *J. Econ. Theory* 11 (1975), 329–339.

JOURNAL OF ECONOMIC THEORY 24, 310 (1981)

Papers to Appear in Forthcoming Issues

The following articles are among those newly accepted for publication in the *Journal of Economic Theory.* They are listed alphabetically by first author.

MICHAEL C. BLAD AND CHRISTOPHER ZEEMAN, Oscilations between Repressed Inflation and Keynesian Equilibria Due to Inertia in Decision Making. Correspondent: Professor Michael C. Blad, Institute of Economics, University of Copenhagen, Studiestraede 6, Copenhagen K DK-1455, Denmark.

LARRY G. EPSTEIN, Comparative Dynamics in the Adjustment-Cost Model of the Firm. Correspondent: Professor Larry G. Epstein, Institute for Policy Analysis, University of Toronto, 150 St. George Street, Toronto M5S 1A1, Canada.

JAMES HESS, Risk and the Gain from Information. Correspondent: Professor James Hess, Department of Economics, University of Southern California, Los Angeles, CA 90002.

RICHARD JENSEN, Adoption and Diffusion of an Innovation of Uncertain Profitability. Correspondent: Professor Richard Jensen, Department of Economics, Ohio State University, Columbus OH, 43210.

ROBERT W. ROSENTHAL, A Dynamic Model of Duopoly with Customer Loyalties. Correspondent: Professor Robert W. Rosenthal, Room 2C-121, Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ. 07974.

AVNER SHAKED AND JOHN SUTTON, Imperfect Information, Perceived Quality, and the Formation of Professional Groups. Correspondent: John Sutton, London School of Economics, Houghton Street, London WC2A 2AF, England.

Printed in Belgium

Copyright © 1981 by Academic Press, Inc.

FORECASTING IN BUSINESS AND ECONOMICS

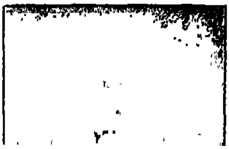
By C. W. J. GRANGER
University of California, San Diego

Written by the recognized authority on forecasting theory and application, *Forecasting in Business and Economics* gives the senior, MBA, and beginning graduate student a clear-cut strategy for tackling forecasting problems

Unlike any other book in the field, this new text stresses the importance of selecting a relevant information set and methods of evaluating the forecast once it has been prepared Granger's emphasis on critical assessment gives the future economist and manager a sound base for interpreting forecasts and making decisions about the future

Modern techniques of statistical, econometric, population, and technological forecasting are presented along with discussions of a wide range of topics including Box-Jenkins methods, leading indicators, survey data, world models, and the use of information sets of increasing size

1980 226 pages



FUNDAMENTALS OF BUSINESS STATISTICS

JEROME D BRAVERMAN, Rider College

A business statistics book that the job-oriented student can really profit from! For each statistical technique introduced, an example of its application is given with a step-by-step solution Students gain a true perspective of how statistical techniques are used in the business world while they gain practice in intelligent decision making

A special *review chapter* early in the book gives students with weak backgrounds the confidence they'll need to succeed in the course Only a knowledge of high school math is required

A *Student Solutions Manual* is available

1979 597 pages

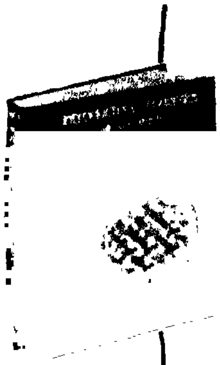
ACADEMIC PRESS, INC.

A Subsidiary of Harcourt Brace Jovanovich, Publishers

NEW YORK LONDON TORONTO SYDNEY SAN FRANCISCO

For examination copies write to College Department ACADEMIC PRESS, INC 111 Fifth Avenue, New York, N Y 10003 Please state course title enrollment, and present text

AP 7964



Just Published . . .

INTRODUCTORY ECONOMICS

by MICHAEL VESETH **University of Puget Sound**

1981 800 pages

The *first* principles text with a supply-side emphasis, this thoroughly lucid incentive approach to the introductory course brings *today's* economics into your classroom Veseth concentrates on the nature of economic problems and the trade-offs inherent in any solution to these problems

The text is available in a *one-volume clothbound edition* or two paperback volumes

THE TEACHING PACKAGE

COURSEBOOK FOR INTRODUCTORY ECONOMICS

by MICHAEL VESETH

1981 300 pages Paperbound

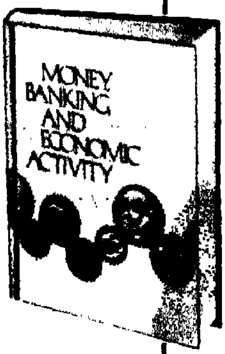
Informative, up-to-date articles demonstrate the workings of economic ideas while a full assortment of questions tests your students' mastery of each chapter.

TEST BANK

by JOHN MARCIS, Kansas State University, and MICHAEL VESETH

2000 multiple-choice questions, *available in book or computer tape formats.*

Instructor's Manual and **Transparency Masters,** gratis upon adoption.



MONEY, BANKING, AND ECONOMIC ACTIVITY

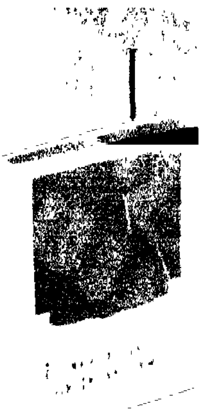
by GAIL E MAKINEN Principal Macroeconomist, United States General Accounting Office, and Georgetown University

1981 556 pages

From one of the nation's leading macroeconomists comes this up-to-date, accurate presentation of the principles of money, banking, and economic activity.

Money, Banking, and Economic Activity incorporates the 1980 changes in the laws governing financial institutions, and is the first text to apply *micro*theory to the analysis of the role of money in our economic system.

Address requests for examination copies to
College Department, ACADEMIC PRESS, INC.,
111 Fifth Avenue, New York, N Y 10003
Please state course title, enrollment, and present text



ECONOMICS:
Private and Public Choice, *Second Edition*
by JAMES D. GWARTNEY, Florida State University
and RICHARD STROUP Montana State University
with the assistance of A H Studenmund,
Occidental College

Rapidly becoming one of the leading principles texts in the field, the second edition shows how modern economic analysis is capable of explaining the turbulence of the last decade and of suggesting policy for dealing effectively with unemployment inflation and sluggish growth More than any other text published in recent years, *ECONOMICS· Private and Public Choice* bridges the gap between traditional economic analysis and the advances that have revolutionized economic thinking

The text is available in a one-volume clothbound edition or two paperback volumes

THE TEACHING PACKAGE

THE COURSEBOOK for ECONOMICS: Private and Public Choice, *Second Edition*
by STROUP/STUDENMUND/GWARTNEY

More than just a study guide this Coursebook contains numerous test questions; problems for detailed analysis and projects, plus a series of current readings

THE TEST BANK
1800 exam items, available in book or computer tape formats

Instructor's Manual and **Transparency Masters,** gratis upon adoption

Other fine publications-

INTERMEDIATE MICROECONOMICS WITH APPLICATIONS. 1980.
AROOP K. MAHANTY

ENERGY ECONOMICS AND POLICY. 1980.
JAMES M GRIFFIN and HENRY B STEELE

FORECASTING IN BUSINESS AND ECONOMICS. 1980.
C.W. J. GRANGER.

MONEY AND INFLATION: A Monetarist Approach. 1975.
J. HUSTON McCULLOCH

Academic Press, Inc.
A Subsidiary of Harcourt Brace Jovanovich, Publishers
NEW YORK • LONDON • TORONTO • SYDNEY • SAN FRANCISCO

APC 1005

How To Comply With The New Copyright Law

Libraries everywhere have found the easy way to fill photocopy requests legally and instantly, without the need to seek permissions, from this and over 3000 other key publications in business, science, humanities, and social science

Participation in the Copyright Clearance Center (CCC) assures you of legal photocopying at the moment of need. You can

Fill requests for multiple copies, interlibrary loan (beyond the CONTU guidelines), and reserve desk without fear of copyright infringement

Supply copies simply and easily from registered publications. The CCC's flexible reporting system accepts photocopying reports and returns an itemized invoice *You need not keep any records*, our computer will do it for you.

The Copyright Clearance Center is your one-stop place for on-the-spot clearance to photocopy for internal use. You will never have to decline a photocopy request or wonder about compliance with the law for any publication registered with the CCC.

For more information, just contact:

Copyright Clearance Center
21 Congress Street
Salem, Massachusetts 01970
(617) 744-3350

a not-for-profit corporation

NAME TITLE

ORGANIZATION

ADDRESS

CITY STATE ZIP

COUNTRY TELEPHONE

JOURNAL OF ECONOMIC THEORY 24, 311–327 (1981)

A Bilateral Monopoly and the Nash Cooperative Solution*

MAMORU KANEKO

Institute of Socio-Economic Planning, University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki-ken 305, Japan

Received June 14, 1978; revised December 12, 1979

1. INTRODUCTION AND THE BILATERAL ECONOMY

Let us consider a bilateral exchange economy, which consists of a seller S and a buyer B. The seller owns $Q > 0$ amount of a commodity, which he wants to sell to the buyer. The buyer owns no amount of the commodity but owns $M > 0$ amount of another commodity, which we call "money." "Money" should be interpreted as the composite commodity of all commodities other than the commodity in question. Of course, the seller may own a positive amount of money, but since we will consider only the case where the seller never buys any amount of the commodity from the buyer, it is convenient to represent the initial level of money by zero. That is, the initial endowments of the seller and buyer are represented as $w^s = (Q, 0)$ and $w^b = (0, M)$. The seller and buyer have von Neumann–Morgenstern utility functions $U(x_1, x_2)$ and $V(x_1, x_2)$ defined on the non-negative orthant E^2_+ of the two-dimensional Euclidean space.

Nash [1, 3] provided a solution concept for two-person cooperative games. If the theory is applied to the above bilateral economy, it selects an allocation $((Q-q, p), (q, M-p))$ such that

$$\begin{aligned}(U(Q-q, p) - U(w^s))(V(q, M-p) - V(w^b))\\ = \max(U(x_1, x_2) - U(w^s))(V(y_1, y_2) - V(w^b))\\ \text{subject to } (x_1, x_2) + (y_1, y_2) = (Q, M), (x_1, x_2), (y_1, y_2) \in E^2_+\\ \text{and } U(x_1, x_2) \geqslant U(w^s), V(y_1, y_2) \geqslant V(w^b). \qquad (1.1)\end{aligned}$$

We call an allocation satisfying (1.1) a *Nash allocation*.

Nash provided two complementary approaches to the solution. One is the axiomatic approach, which was well established in Nash [1]. The axiomatic approach focuses its consideration only on outcomes of a bargaining process instead of the bargaining process itself. Nash [3] provided another approach

---

* The research in this paper was partially sponsored by the office of Naval Research Contract Number N00014-77-0518.

0022-0531/81/030311-17$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

to it in which the solution was characterized as a kind of equilibrium point of a game in extensive form. This approach intends to exhibit a structure of bargaining process yielding the Nash coorperative solution. As the latter approach provides an explicit structure of bargaining process, it looks more realistic than the former. But when this approach is applied to the above bilateral economy, the bargaining process never seems to fit the bilateral economy, i.e., the extensive game given by Nash cannot be naturally interpreted in the context of the bilateral economy. His extensive game should be thought of just as an archetype. The purpose of this paper is to modify the latter approach to incorporate the case of the bilateral economy.

Without loss of generality we can put $U(w^s) = V(w^b) = 0$. We assume the following conditions on the utility functions U and V:

(A) U and V are monotonically increasing, i.e., if $x \geqslant y$ and $x \neq y$, then $U(x) > U(y)$ and $V(x) > V(y)$;

(B) U and V are functions of C^2, i.e., U and V have all second derivatives which are continuous;

(C) U and V are strictly concave functions;

(D) $U(0, M)\, V(Q, 0) < 0$, but there is an m such that $U(0, m)$, $V(Q, M - m) > 0$;

(E) for each m with $U(0, m)\, V(Q, M - m) > 0$, there is a (q, m') such that $U(Q - q, m')\, V(q, M - m') > U(0, m)\, V(Q, M - m)$.

Assumptions (A), (C), (D) and (E) impose regularities upon the bilateral economy. Assumption (B) is a mathematical condition. Assumptions (D) and (E) are not familiar to us, but they just claim that the Nash allocation does not lies on the boundary.

Before we provide the last assumption, we need to state the following lemmas, which will be proved in the Appendix.

LEMMA 1. *For each* $q \in (0, Q]$ *there is an* $m \in (0, M]$ *such that* $U(Q - q, m)$, $V(q, M - m) > 0$, *but there is no* m *such that* $U(Q - q, m)$, $V(q, M - m) \leqslant 0$.

LEMMA 2. *For each* $q \in (0, Q]$, *there is a unique* $m(q) \in (0, M)$ *such that*

$$U(Q - q, m(q))\, V(q, M - m(q)) \geqslant U(Q - q, m)\, V(q, M - m) \quad \text{for all} \quad m \in (0, M]. \tag{1.2}$$

LEMMA 3. *Let* $u(q) = U(Q - q, m(q))$, $v(q) = V(q, M - m(q))$ *and* $L(q) = \log u(q) + \log v(q)$ *for all* $q \in (0, Q]$. *Then* $L(q)$ *is a strictly concave function.*

It should be noted that $u(q) > 0$ and $v(q) > 0$ for all $q \in (0, Q]$ by Lemmas 1 and 2.

LEMMA 4. (i) *There is a unique* $q_N \in (0, Q)$ *such that* $L(q_N) \geqslant L(q)$ *for all* $q \in (0, Q]$.

(ii) $((Q - q_N, m(q_N)), (q_N, M - m(q_N))$ *is a unique Nash allocation.*

The last assumption is as follows:

(F) $\log u(q)$ *and* $\log v(q)$ *are concave functions of* q *on* $(0, Q]$.

Since, as shown in Lemma 3, it follows from (A)–(E) that $\log u(q) + \log v(q)$ is a strictly concave function, Assumption (F) would be a weak condition. But the author has succeeded in neither proving (F) from (A)–(E) nor constructing any counterexample. The following example satisfies all the assumptions.

EXAMPLE 1. Let $U(x_1, x_2) = x_1^{1/2} + x_2 - 10$, $V(x_1, x_2) = \frac{3}{2} \cdot x_1^{1/2} + x_2 - 20$, $Q = 100$ and $M = 20$. It is not difficult to verify that Assumptions (A)–(E) are satisfied. Then the program

$$\max_{0 < m \leqslant M} \log(\sqrt{100 - q} + m - 10) + \log(\tfrac{3}{2}\sqrt{q} - m)$$

has the solution $m(q) = (\frac{3}{2}\sqrt{q} - \sqrt{100 - q} + 10)/2$ for each $q \in (0, Q]$. Hence $u(q)$ and $v(q)$ are represented as

$$u(q) = v(q) = (\tfrac{3}{2}\sqrt{q} + \sqrt{100 - q} - 10)/2.$$

It is clear that this function itself is a concave function with $u(q) = v(q) > 0$ for all $q \in (0, Q]$.

2. THE BARGAINING GAME $G(f_0, g_0)$

The seller and buyer have a negotiation to abide by the rules of the bargaining game $G(f_0, g_0)$. We assume that the bargaining game $G(f_0, g)$ is divided into two stages 1 and 2. At the 1st stage, an amount of the commodity to be traded is decided, and at the 2nd stage, an amount of money to be paid to the seller from the buyer is decided. The rules of the bargaining game are as follows:

The 1*st stage.* The seller S and buyer B select amounts of the commodity that they want to trade from S to B. These decisions are made independently. That is, each must select an amount without the knowledge of the decision of the other. Then they inform each other of the amounts *selected. If the decisions are the same,* then the **nogetiation goes on to the** *2nd stage, but, otherwise, the negotiation* is broken off.

The strategy-spaces of S and B at the 1st stage are (0, Q]. It should be noted that 0 does not belong to the strategy-space, since it is assumed that they want to trade the commodity. Let $f_0(q_s - q_b)$ *be the function on* $(-Q, Q)$ defined by

$$f_0(q_s - q_b) = 1 \quad \text{if} \quad q_s - q_b = 0 \qquad\qquad (2.1)$$
$$= 0 \quad \text{otherwise,}$$

where q_s and q_b are amounts of the commodity selected by S and B, respectively. The rule can be rewritten by this notation as follows. If S and B select strategies q_s and q_b, then the negoriation goes on to the 2nd stage with probability $f_0(q_s - q_b)$ and is broken off with probability $1 - f_0(q_s - q_b)$.

The 2nd stage. Let the negotiation go on to the 2nd stage. This case happens if and only if $q_s = q_b = \bar{q}$. But if S and B violate the rule and if they make the negotiation go on to the 2nd stage and be concluded in a case of $q_s \neq q_b$, then it is assumed that $(q_s + q_b)/2 = \bar{q}$ amount of the commodity is traded. At this stage, S and B select prices of the amount $\bar{q}$ of the commodity. These decisions are also made independently, and they inform each other of the prices. The negotiation is concluded if and only if price p_s set by S coincides with price p_b set by B. Similar to the 1st stage, we take the violation of this rule into account and assume that if they make the negotiation be concluded in a case of $p_s \neq p_b$, then the final price is $\bar{p} = (p_s + p_b)/2$. It should be noted that p_s, p_b or $\bar{p}$ are the total amount of money to be paid from B to S, and that the unit prices are given as $p_s/\bar{q}$, $p_b/\bar{q}$ or $\bar{p}/\bar{q}$, respectively. Let g_0 be the function defined by

$$g_0(p_s - p_b) = 1 \quad \text{if} \quad p_s - p_b = 0 \qquad\qquad (2.2)$$
$$= 0 \quad \text{otherwise.}$$

The set of prices to be set by S or B is $(0, M]$. Hence the domain of g_0 is the open interval $(-M, M)$. 0-price is omitted from the set of prices by the same reason as that of the omission of 0-quantity. The above rule can be also rewritten by g_0 as follows: The negotiation is concluded with probability $g_0(p_s - p_b)$ and is broken off with probability $1 - g_0(p_s - p_b)$.

The bargaining game $G(f_0, g_0)$ is formulated as a game in extensive form. The game tree of the game is drawn as Fig. 1. Of course, the order of S's and B's information sets is exchangeable. Even if $q_s \neq q_b$, then the game tree is assumed to have the 2nd stage, though the probability that the negotiation goes on to the 2nd stage is zero. A *behavior strategy* of S is a pair (q_s, t_s) of a point q_s in $(0, Q]$ and a function t_s from $(0, Q]^2 \equiv (0, Q] \times (0, Q]$ to $(0, M]$. A *behavior strategy* of B is also a pair (q_b, t_b) of a point q_b in $(0, Q]$

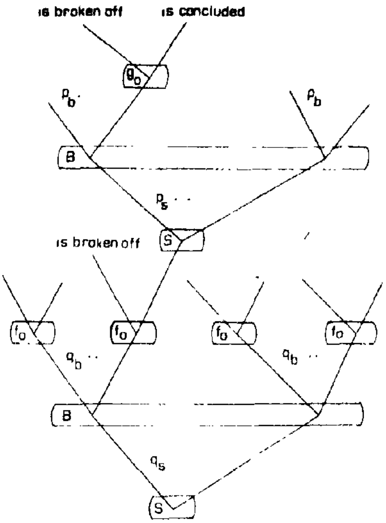


FIGURE 1

and a function t_b from $(0, Q]^2$ to $(0, M]$. The set of all behavior strategies of S is the same as that of all behavior strategies of B, which is denoted by $BS \equiv (0, Q] \times T$. The *payoff functions* K_s and K_b of the bargaining game $G(f_0, g_0)$ are real valued functions on $BS^2 = BS \times BS$. If S and B use behavior strategies $b_s = (q_s, t_s)$ and $b_b = (q_b, t_b)$, respectively, the payoff functions are represented as

$$\begin{aligned} K_s(b_s, b_b) &= U(Q - \bar{q}, \bar{p}) f_0(q_s - q_b) g_0(p_s - p_b), \\ K_b(b_s, b_b) &= V(\bar{q}, M - \bar{p}) f_0(q_s - q_b) g_0(p_s - p_b), \end{aligned} \tag{2.3}$$

where $p_s = t_s(q_s, q_b)$, $p_b = t_b(q_s, q_b)$, $\bar{q} = (q_s + q_b)/2$ and $\bar{p} = (p_s + p_b)/2$.

The 2nd stage of the bargaining game $G(f_0, g_0)$ is subgames of $G(f_0, g_0)$. For definition of subgame, see Selten [4, 5]. When q_s and q_b are selected at the 1st stage, the subgame of the 2nd stage is denoted by $SG(q_s, q_b, g_0)$. $SG(q_s, q_b, g_0)$ is characterized as a game as follows. The set of strategies of S or B is $(0, M]$, which is the set of values of behavior strategies $t_s(q_s, q_b)$ or $t_b(q_s, q_b)$, i.e., $(0, M] = \{t(q_s, q_b) \mid t \in T\}$. The payoff functions k_s and k_b of $SG(q_s, q_b, g_0)$ are given as

$$\begin{aligned} k_s(p_s, p_b) &= U(Q - \bar{q}, \bar{p}) g_0(p_s - p_b), \\ k_b(p_s, p_b) &= V(\bar{q}, M - \bar{p}) g_0(p_s - p_b), \end{aligned} \tag{2.4}$$

where $\bar{q} = (q_s + q_b)/2$ and $\bar{p} = (p_s + p_b)/2$. We call k_s and k_b the *induced payoff functions* of K_s and K_b on $SG(q_s, q_q, g_0)$. We call $p_s = t_s(q_s, q_b)$ and

$p_b = t_b(q_s, q_b)$ the *induced strategies* of $b_s = (q_s, t_s)$ and $b_b = (q_b, t_b)$ on $SG(q_s, q_b, g_0)$.

When strategies t_s and t_b in T at the 2nd stage are fixed, $G(f_0, g_0)$ is reduced to a game with one stage. We call the game a *reduced game*, which is denoted by $RG(f_0, g_0, t_s, t_b)$. It is characterized as follows. The sets of all strategies are $(0, Q]$, and S's and B's payoff functions r_s and r_b of $RG(f_0, g_0, t_s, t_b)$ are represented as

$$\begin{aligned} r_s(q_s, q_b) &= K_s((q_s, t_s), (q_b, t_b)), \\ r_b(q_s, q_b) &= K_b((q_s, t_s), (q_b, t_b)). \end{aligned} \tag{2.5}$$

Let (X_1, X_2, k_1, k_2) be a two-person game, where X_1 and X_2 are the sets of strategies, and where k_1 and k_2 are the payoff functions. Then we call a pair $(x_1^*, x_2^*) \in X_1 \times X_2$ of strategies an equilibrium point if

$$\begin{aligned} k_1(x_1^*, x_2^*) &\geqslant k_1(x_1, x_2^*) \qquad \text{for all} \quad x_1 \in X_1, \\ k_2(x_1^*, x_2^*) &\geqslant k_2(x_1^*, x_2) \qquad \text{for all} \quad x_2 \in X_2. \end{aligned} \tag{2.6}$$

We call a pair $(b_s^*, b_b^*) = ((q_s^*, t_s^*), (q_b^*, t_b^*))$ of behavior strategies of $G(f_0, g_0)$ an *subgame perfect equilibrium point* if (b_s^*, b_b^*) satisfies the following (2.7) and (2.8);

For each $(q_s, q_b) \in (0, Q]^2$, the pair (p_s, p_b) of the induced strategies of b_s^* and b_b^* on $SG(q_s, q_b, g_0)$ is an equilibrium point of $SG(q_s, q_b, g_0)$. (2.7)

(q_s^*, q_b^*) is an equilibrium point of the reduced game $RG(f_0, g_0, t_s^*, t_b^*)$. (2.8)

This solution concept is defined by Selten, and the basic idea was explained in Selten [4, 5].

In the bargaining game $G(f_0, g_0)$ there are a lot of subgame perfect equilibrium points. For example, if a pair $(b_s^*, b_b^*) = ((q_s^*, t_s^*), (q_b^*, t_b^*))$ satisfies

$$q_s^* = q_b^* \qquad \text{and} \qquad t_s^*(q_s, q_b) = t_b^*(q_s, q_b) \qquad \text{for all} \quad (q_s, q_b) \in (0, Q]^2, \tag{2.9}$$

$$U(Q - (q_s + q_b)/2, t_s^*(q_s, q_b)) \geqslant 0 \quad \text{and} \quad V((q_s + q_b)/2, M - t_b^*(q_s, q_b)) \geqslant 0 \qquad \text{for all} \quad (q_s, q_b) \in (0, Q]^2, \tag{2.10}$$

then it is a subgame perfect equilibrium point. Hence we would like to select an equilibrium point from the set of these equilibrium points by comparing relative stabilities of these points. In the next section, we shall define a solution concept for this purpose.

3. THE PERFECT EQUILIBRIUM POINT OF $G(f_0, g_0)$

We shall define a solution concept being based on the idea of the smoothing procedure introduced by Nash [3]. In our case, the basic idea is that the bargaining game $G(f_0, g_0)$ is thought of as a limit case of bargaining games in which the probabilities that S and B violate the rules of the 1st and 2nd stages, i.e., f_0 and g_0, take positive values. Selten [5] reconsidered the concept of his subgame perfect equilibrium point in extensive games and defined a solution concept called "perfect equilibrium point" due to the similar procedure. The approach of this paper has the same thought with Selten's.

We shall approximate f_0 and g_0 by classes of smooth functions. Let h_0 be the function on $(-K, K)$ such that

$$\begin{aligned} h_0(w) &= 1 \qquad \text{if} \quad w = 0 \\ &= 0 \qquad \text{otherwise,} \end{aligned}$$

where K is a positive constant. Of course, when $K = Q$ or $K = M$, $h_0 = f_0$ or $h_0 = g_0$, respectively. We call $\{h_n\}$ an *approximate sequence* of h_0 if

for all $n \geqslant 1$, $h_n(w)$ is a continuously differentiable function on $(-K, K)$ with $0 < h_n(w) \leqslant 1$ and $h_n(0) = 1$, (3.1)

$\log h_n(w)$ is a strictly concave function for all $n \geqslant 1$, (3.2)

$\{h'_n(w)/h_n(w)\}$ converges pointwise to $+\infty$ on $(-K, 0)$ and to $-\infty$ on $(0, K)$. (3.3)

Here $h'_n = dh_n/dw$. It is not difficult to verify that these conditions imply

$$\lim_{n \to \infty} h_n(w) = h_0(w) \qquad \text{for each } w \in (-K, K), \tag{3.4}$$

$\log h_n(w)$, i.e., $h_n(w)$ itself is increasing on $(-K, 0)$ and decreasing on $(0, K)$ for each $n \geqslant 1$, (3.5)

$\{h'_n/h_n\}$ converges uniformly to $+\infty$ on $(-K, -\varepsilon)$ and to $-\infty$ on (ε, K) for any $\varepsilon > 0$. (3.6)

For example, it is the case if $h_n(w) = e^{-nw^2}$.

Let $\{f_n\}$ and $\{g_n\}$ be approximate sequences of f_0 and g_0, respectively. In

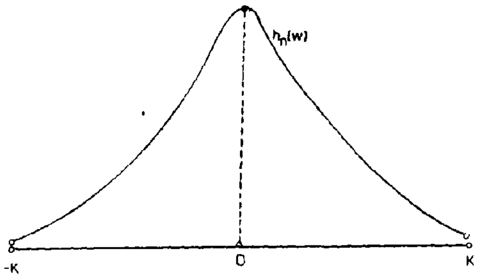


FIGURE 2

the following, we shall consider the games $SG(q_s, q_b, g_n)$ and $RG(f_n, g_0, t_s, t_b)$ in which f_0 and g_0 are substituted by f_n and g_n. The definitions of games $SG(q_s, q_b, g_n)$ and $RG(f_n, g_0, t_s, t_b)$ are clear and unmistakable. We shall think of $SG(q_s, q_b, g_0)$ and $RG(f_0, g_0, t_s, t_b)$ as limit cases of $\{SG(q_s, q_b, g_n)\}$ and $\{RG(f_n, g_0, t_s, t_b)\}$.

A function f_n assigns the probability that the negotiation goes on to the 2nd stage. When q_s is not equal to q_b, $f_n(q_s - q_b)$ is the probability that S and B violate the rule of the 1st stage and make the negotiation go on to the 2nd stage. Similarly, g_n assigns the probability that the negotiation is concluded, assuming that it goes on to the 2nd stage. That is, when $p_s \neq p_b$, $g_n(p_s - p_b)$ is the probability that S and B violate the rule of the 2nd stage and make the negotiation be concluded. Condition (3.1) means that the probability of the violation is always positive, but (3.3) that it becomes quickly smaller as the disparity between q_s and q_b ($p_s$and p_b) becomes larger. Equation (3.2) is a regularity condition.

We call a pair $(b_s^*, b_b^*) = ((q_s^*, t_s^*), (q_b^*, t_b^*))$ of behavior strategies a *perfect equilibrium point* if it satisfies (3.7), (3.8) and (3.9);

(b_s^*, b_b^*) is a subgame perfect equilibrium point of $G(f_0, g_0)$. (3.7)

Let $\{g_n\}$ be an arbitrary approximate sequence of g_0 and let (q_s, q_b) be arbitrary in $(0, Q]^2$. Let $\{E_n\}$ be the sequence such that each E_n is the set of equilibrium points of $SG(q_s, q_b, g_n)$. Then the sequence $\{E_n\}$ converges to $\{(\bar{p}_s, \bar{p}_b)\}$ in the sense of the Hausdorf metric for subsets. Here $(\bar{p}_s, \bar{p}_b)$ is the induced strategy pair of (b_s^*, b_b^*) on $SG(q_s, q_b, g_0)$. (3.8)

Let $\{f_n\}$ be an arbitrary approximate sequence of f_0. Let $\{Q_n\}$ be the sequence such that each Q_n is the set of equilibrium points of $RG(f_n, g_0, t_s^*, t_b^*)$. Then the sequence $\{Q_n\}$ converges to $\{(q_s^*, q_b^*)\}$ in the sense of the Hausdorf metric for subsets. (3.9)

In fact, since E_n and Q_n are singleton sets for sufficiently large n as will be shown in the subsequent section, it is not necessary to use the Hausdorf metric for subsets instead of the usual Euclidean metric in (3.8) and (3.9).

This definition is based on the intuitive idea that a reasonable equilibrium point should be stable against arbitrarily small imperfections of rationality. This paper makes probabilities of violations of or deviations from the rules of game represent imperfections of rationality. Approximate sequences $\{f_n\}$ and $\{g_n\}$ represent arbitrarily small probabilities of violations of or deviations from the rules of game. The above intuitive idea is rewritten in terms of approximate sequences as follows: For any approximate sequences a reasonable equilibrium point in the limit of the sets of all equilibrium points of corresponding games. That is, when a small violation of the rules of game occurs, a reasonable equilibrium point changes a little. This is a continuity property, which is really the requirement of the definition of perfect equilibrium point. It is noted that we impose this requirement upon every subgame $SG(q_s, q_b, g_0)$ and the reduced game $RG(f_0, g_0, t_s^*, t_b^*)$. The reason for this is the same with that for the definition of subgame perfect equilibrium point by Selten. The basic assumption for this is that when a player makes a decision on an information set in a game tree, he uses a rationality to make it supposing that when he will make a decision on a succeeding information set, he will use the same rationality again. This would be a natural assumption. This is also the reason why only f_0 is substituted by f_n but not g_0 in $RG(f_n, g_0, t_s^*, t_b^*)$ of (3.9). In sum, we require perfect equilibrium point to have the stability against small imperfections of rationality on every information set.

Our main result of this paper is the following theorem:

MAIN THEOREM. *There exists a unique perfect equilibrium point of the bargaining game $G(f_0, g_0)$. The perfect equilibrium point $(b_s^*, b_b^*) = ((q_s^*, t_s^*), (q_b^*, t_b^*))$ is represented as*

$$q_s^* = q_b^* = q_N \qquad \textit{and} \qquad t_s^*(q_s, q_b) = t_b^*(q_s, q_b) = m((q_s + q_b)/2)$$
$$\textit{for all} \quad (q_s, q_b) \in (0, Q]^2, \quad (3.10)$$

where q_N and $m(q)$ are given in Lemmas 4 and 2, respectively.

This theorem means that the perfect equilibrium point yields the Nash allocation $((Q - q_N, m(q_N)), (q_N, M - m(q_N)))$, and that even if $(q_s, q_b) \neq (q_N, q_N)$ are played but if the negotiation goes on to the 2nd stage, the perfect equilibrium point also yields the Nash cooperative solution $((Q - (q_s + q_b)/2, m((q_s + q_b)/2)), ((q_s + q_b)/2, M - m((q_s + q_b)/2)))$ in the subgame of the 2nd stage.

EXAMPLE 2. Let us consider the bilateral economy given in Example 1. Then the perfect equilibrium point $(b_s^*, b_b^*) = ((q_s^*, t_s^*), (q_b^*, t_b^*))$ is given as

$$\begin{aligned} q_s^* = q_b^* &= q_N = 900/13, \\ t_s^*(q_s, q_b) = t_b^*(q_s, q_b) &= m((q_s + q_b)/2) \\ &= (\tfrac{3}{2}\sqrt{(q_s + q_b)/2} - \sqrt{100 - (q_s + q_b)/2} + 10)/2 \\ &\text{for all} \quad (q_s, q_b) \in (0, Q]^2. \end{aligned}$$

The Nash allocation is $((Q - q_N, m(q_N)), (q_N)), (q_N, M - m(q_N)) = ((400/13, (25 + 10\sqrt{13})/2\sqrt{13}), (900/13, (30\sqrt{13} - 25)/2\sqrt{13})) \doteqdot ((30.8, 8.5), (69.2, 11.5))$.

We have shown that the Nash coorperative solution is derived as the perfect equilibrium point in the bargaining game $G(f_0, g_0)$. The bargaining game $G(f_0, g_0)$, however, is one of several models which we can consider. For example, we can formulate other models of the bilateral monopoly, in which a quantity of the 1st commodity and a price of it are decided simultaneously, or in which a price of it is decided before a quantity is done, etc. In such models our basic approach of smoothing procedure may also choose a narrower class of equilibrium points from those of an original game. It should be, however, noted that the Nash allocation is not necessarily derived. The reason for the variety of models of the bilateral monopoly is that it has two decision-variables, i.e., it has the two kinds of commodities. For example, when an amount of the 1st commodity to be traded is fixed by an exogeneous decision, or when the commodity has a very large and indivisible unit, e.g., a house, a tanker, a petrochemical plant, etc., the game of the 2nd stage would be relevant to such an economy. It may be, however, more natural to assume instead of (2.2) in the 2nd stage:

$$\begin{aligned} g_0(p_s - p_b) &= 1 \qquad \text{if} \quad p_s - p_b \leqslant 0 \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

In this case, our theorem still hold. Thus it is thought that our basic approach may be applied to wide fields.

4. PROOF OF MAIN THEOREM

Initially, apart from the bargaining game $G(f_0, g_0)$, we prepare three auxiliary theorems which will be used in the proof of the Main Theorem.

Let us consider a sequence of games $\{J_n\} = \{(X_1, X_2, k_1^n, k_2^n)\}$ such that for all $n \geqslant 1$,

$$\begin{aligned} X_1 &= X_2 = (0, K], \\ k_1(x_1, x_2) &= u_1((x_1 + x_2)/2)\, h_n(x_1 - x_2) \\ k_2(x_1, x_2) &= u_2((x_1 + x_2)/2)\, h_n(x_1 - x_2) \\ &\text{for all} \quad (x_1, x_2) \in (0, K]^2, \end{aligned} \tag{4.1}$$

where $u_1(z)$ and $u_2(z)$ are real-valued functions on $(0, K]$ and $\{h_n\}$ is an approximate sequence of h_0.

Auxiliary Theorem A. *Assume*:

(A_1) u_1 *and* u_2 *are differentiable functions of* z.

(A_2) $\log u_1$ *and* $\log u_2$ *are concave functions on* $\{z \in (0, K]: u_1(z) > 0\}$ *and* $\{z \in (0, K]: u_2(z) > 0\}$, *respectively.*

(A_3) *There is a unique* $z^* \in (0, K)$ *such that* $\log u_1(z^*) + \log u_2(z^*) \geqslant \log u_1(z) + \log u_2(z)$ *for all* $z \in \{z \in (0, K]: u_1(z), u_2(z) > 0\}$.

(A_4) *There is an integer* N *such that for all* $n \geqslant N$, *there exists a unique equilibrium point* (x_1^n, x_2^n) *of* J_n *with* $x_1^n, x_2^n \in (0, K)$, $u_1((x_1^n + x_2^n)/2) > 0$ *and* $u_2((x_1^n + x_2^n)/2) > 0$.

Then it holds that $(x_1^n + x_2^n)/2 = z^*$ *for all* $n \geqslant N$ *and* $\lim_{n \to \infty} x_1^n = \lim_{n \to \infty} x_2^n = z^*$.

Proof. It follows from (A_1), (A_2) and (A_4) that for each $n \geqslant N$, (x_1^n, x_2^n) is an equilibrium point of J_n if and only if $(\partial/\partial x_1) \log k_1^n = (\partial/\partial x_2) \log k_2^n = 0$, that is,

$$\frac{u_1'}{u_1} = -\frac{u_2'}{u_2} = -2 \cdot \frac{h_n'}{h_n}. \tag{4.2}$$

It also follows from (A_2) and (A_3) that z coincides with z^* if and only if $(d/dz)(\log u_1(z) + \log u_2(z))|_{z=z^*} = 0$, that is,

$$\frac{u_1'(z^*)}{u_1(z^*)} = -\frac{u_2'(z^*)}{u_2(z^*)}. \tag{4.3}$$

Since there is a unique z^* satisfying (4.3) by (A_3), we have $z^* = (x_1^n + x_2^n)/2$ for each $n \geqslant N$. If $\{x_1^n - x_2^n\}$ does not converge to 0, then there are an $\varepsilon > 0$ and a subsequence $\{x_1^{n''} - x_2^{n''}\}$ such that $|x_1^{n''} - x_2^{n''}| \geqslant \varepsilon$ for all n''. Hence $\{h_{n''}'(x_1^{n''} - x_2^{n''})/h_{n''}(x_1^{n''} - x_2^{n''})\}$ converges uniformly to $+\infty$ or $-\infty$ by (3.6). This contradicts (4.2). Hence $\{x_1^n - x_2^n\}$ must converge to 0. Hence we have $\lim_{n \to \infty} x_1^n = \lim_{n \to \infty} x_2^n = z^*$. Q.E.D.

AUXILIARY THEOREM B. *Assume*:

(B$_1$) u_1 *and* u_2 *are differentiable and concave functions of* z.

(B$_2$) u_1 *and* u_2 *are increasing and decreasing on* $(0, K]$, *respectively*.

(B$_3$) *There is a* $\bar{z} \in (0, K)$ *such that* $u_1(\bar{z}) > 0$ *and* $u_2(\bar{z}) > 0$.

(B$_4$) $\inf u_1(z) < 0$ *and* $u_2(K) < 0$.

Then Assumptions (A$_1$), (A$_2$), (A$_3$) *and* (A$_4$) *of Auxiliary Theorem* A *hold*.

Proof. It is easy to verify that (B$_1$)–(B$_4$) imply (A$_1$), (A$_2$) and (A$_3$). We will show (A$_4$). To do so, we need to prove two claims.

Claim 1. If $u_1((x_1 + x_2)/2) \leqslant 0$ or $u_2((x_1 + x_2)/2) \leqslant 0$, then there is an integer N_1 such that (x_1, x_2) can not be an equilibrium point of J_n for all $n \geqslant N_1$.

It follows from (B$_2$) and (B$_3$) that $u_1((x_1 + x_2)/2) \leqslant 0$ implies $u_2((x_1 + x_2)/2) > 0$ and $u_2((x_1 + x_2)/2) \leqslant 0$ implies $u_1((x + x_2)/2) > 0$. Let us consider the first case, i.e., $u_1((x_1 + x_2)/2) \leqslant 0$ and $u_2((x_1 + x_2)/2) > 0$. Let ε be a sufficiently small positive number. If $|x_1 - x_2| \geqslant \varepsilon$, then there is an integer N_1 by (3.6) such that for all $n \geqslant N_1$,

$$\frac{u_2'((x_1 + x_2)/2)}{u_2(x_1 + x_2)/2)} > 2\frac{h_n'(x_1 - x_2)}{h_n(x_1 - x_2)} \qquad \text{if} \quad x_1 > x_2,$$

$$\frac{u_2'((x_1 + x_2)/2)}{u_2((x_1 + x_2)/2)} < 2\frac{h_n'(x_1 - x_2)}{h_n(x_1 - x_2)} \qquad \text{if} \quad x_1 < x_2.$$

That is,

$$\frac{\partial}{\partial x_2} \log k_2^n(x_1, x_2) > 0 \qquad \text{if} \quad x_1 > x_2,$$

$$\frac{\partial}{\partial x_2} \log k_2^n(x_1, x_2) < 0 \qquad \text{if} \quad x_1 < x_2.$$

This means that (x_1, x_2) can not be an equilibrium point of J_n for all $n \geqslant N_1$. Suppose $|x_1 - x_2| < \varepsilon$. If $x_1 = K$ or $x_2 = K$, then $u_2((x_1 + x_2)/2) < 0$ by (B$_4$), because ε is sufficiently small, which is a contradiction. Hence we can assume $x_1, x_2 \in (0, K)$. If $u_1((x_1 + x_2)/2) = 0$, then it holds by (B$_2$) that for $y_1 > x_1$,

$$k_1^n(y_1, x_2) = u_1((y_1 + x_2)/2)\, h_n(y_1 - x_2) > 0 = k_1^n(x_1, x_2),$$

which means that (x_1, x_2) can not be an equilibrium point of J_n for all $n \geqslant 1$.

Let $u_1(x_1, x_2) < 0$. If $x_1 \leqslant x_2$, then by (3.5), $h_n'(x_1 - x_2)/h_n(x_1 - x_2) \geqslant 0$ for all $n \geqslant 1$. Since $u_2'((x_1 + x_2)/2)/u_2((x_1 + x_2)/2) < 0$ by (B_2), we have

$$\frac{u_2'((x_1 + x_2)/2)}{u_2((x_1 + x_2)/2)} - 2\frac{h_n'(x_1 - x_2)}{h_n(x_1 - x_2)} < 0 \qquad \text{for all} \quad n \geqslant 1,$$

i.e.,

$$\frac{\partial}{\partial x_2} \log k_2^n(x_1, x_2) < 0 \quad \text{for all} \quad n \geqslant 1.$$

This means that (x_1, x_2) cannot be an equilibrium point of J_n for all $n \geqslant 1$. If $x_1 > x_2$, then $\log h_n(y_1 - x_2) < \log h_n(x_1 - x_2)$, i.e., $h_n(y_1 - x_2) < h_n(x_1 - x_2)$ for all $y_1 > x_1$ by (3.5). Further since $u_1((y_1 + x_2)/2) > u_1((x_1 + x_2)/2)$ for all $y_1 > x_1$ by (B_2), we have

$$\begin{aligned} k_1^n(y_1, x_2) &= u_1((y_1 + x_2)/2)h_n(y_1 - x_2) > u_1((x_1 + x_2)/2)h_n(x_1 - x_2) \\ &= k_1^n(x_1, x_2) \qquad \text{for all} \quad y_1 > x_1. \end{aligned}$$

This means that (x_1, x_2) can not be an equilibrium point of J_n for all $n \geqslant 1$.

In the case where $u_1((x_1 + x_2)/2) > 0$ and $u_2((x_1 + x_2)/2) \leqslant 0$, we can analogously prove the claim.

Claim 2. If $x_1 = K$ or $x_2 = K$, then there is an integer N_2 such that (x_1, x_2) cannot be an equilibrium point of J_n for all $n \geqslant N_2$.

We need to consider the case where $u_2((x_1 + x_2)/2) > 0$. Let $x_1 = K$. Since $u_2(K) < 0$ by (B_4), there is an $\bar{x}_2$ such that $u_2((x_1 + \bar{x}_2)/2) = 0$ and $x_2 < \bar{x}_2 < K$. Since $\{h_n'/h_n\}$ converges uniformly to $-\infty$ on $(K - \bar{x}_2, K)$, there is an integer N_2 such that for all $n \geqslant N_2$,

$$\frac{u_2'((x_1 + x_2)/2)}{u_2((x_1 + x_2)/2)} > 2 \cdot \frac{h_n'(x_1 - x_2)}{h_n(x_1 - x_2)},$$

i.e.,

$$\frac{\partial}{\partial x_2} \log k_2^n(x_1, x_2) > 0 \qquad \text{for all} \quad n \geqslant N_2.$$

Since $x_2 < K$, it follows from this that (x_1, x_2) cannot be an equilibrium point of J_n for all $n \geqslant N_2$. In the case where $x_2 = K$, we can analogously prove the claim.

It follows from these claims that for each $n \geqslant N_3 \equiv \max(N_1, N_2)$, a necessary and sufficient condition for (x_1, x_2) to be an equilibrium point of J_n is

$$\frac{\partial}{\partial x_1} \log k_1^n(x_1, x_2) = \frac{\partial}{\partial x_2} \log k_2^n(x_1, x_2) = 0,$$

i.e.,

$$\frac{u_1'((x_1+x_2)/2)}{u_1((x_1+x_2)/2)} = -\frac{u_2'((x_1+x_2)/2)}{u_2((x_1+x_2)/2)} = -2 \cdot \frac{h_n'(x_1-x_2)}{h_n(x_1-x_2)}. \tag{4.4}$$

By (4.4) we have $(x_1+x_2)/2 = z^*$, where z^* is the number given in (A_3). Since for any $\varepsilon > 0$ there is an N_4 by (3.6) such that for all $n \geqslant N_4$,

$$-2 \cdot \frac{h_n'(\varepsilon)}{h_n(\varepsilon)} > \frac{u_1'(z^*)}{u_1(z^*)} > -2 \cdot \frac{h_n'(-\varepsilon)}{h_n(-\varepsilon)}.$$

Since $h_n'(w)/h_n(w)$ is a continuous and decreasing function of w on $(-K, K)$ by (3.1) and (3.2), there is a unique w_n for each $n \geqslant N_4$ such that

$$-2 \cdot \frac{h_n'(w_n)}{h_n(w_n)} = \frac{u_1'(z^*)}{u_1(z^*)}.$$

Hence $(x_1^n, x_2^n) \equiv (z^* + w_n/2,\ z^* - w_n/2)$ satisfies (4.4) for each $n \geqslant \max(N_3, N_4)$, i.e., this (x_1^n, x_2^n) is a unique equilibrium point of J_n for each $n \geqslant N \equiv \max(N_3, N_4)$. Q.E.D.

AUXILIARY THEOREM C. *Assume that* $u_1(z) > 0$ *and* $u_2(z) > 0$ *for all* $z \in (0, K]$ *and that* (A_1), (A_2) *and* (A_3) *are true. Then* (A_4) *is true.*

Proof. Suppose that if $x_1 = K$ or $x_2 = K$, then (x_1, x_2) cannot be an equilibrium point of J_n for all $n \geqslant$ some N. Then it can be proved in the same way with the proof of Auxiliary Theorem B that (A_4) is true. Hence we will show the above supposition.

Since $z^* \in (0, K)$, we have $u_1'(K)/u_1(K) + u_2'(K)/u_2(K) < 0$. Then it must hold that $u_1'(K)/u_1(K) < 0$ or $u_2'(K)/u_2(K) < 0$. Let δ be a positive real number such that $u_1'(K-\delta)/u_1(K-\delta) + u_2'(K-\delta)/u_2(K-\delta) < 0$.

Suppose $x_1 = K$ and $x_2 < K - \delta$. Since $\{h_n'/h_n\}$ converges uniformly to $-\infty$ on (δ, K), there is N such that

$$\frac{u_1'((x_1+x_2)/2)}{u_1((x_1+x_2)/2)} < -2 \cdot \frac{h_n'(x_1-x_2)}{h_n(x_1-x_2)},$$

i.e., $(\partial/\partial x_1) \log k_1^n(x_1, x_2) < 0$. This means that (x_1, x_2) cannot be an equilibrium point of J_n for all $n \geqslant N$. In the case where $x_1 < K - \delta$ and $x_2 = K$, we can prove similarly that (x_1, x_2) cannot be an equilibrium point of J_n for all $n >$ some N.

Suppose that $x_1, x_2 \in [K-\delta, K]$. Let us consider the case where $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) \leqslant u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2) \leqslant 0$. Of course, it holds that $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) < 0$. If $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) < -2h_n'(x_1-x_2)/h_n(x_1-x_2)$, then $(\partial/\partial x_1) \log k_1^n(x_1, x_2) < 0$, which implies that (x_1, x_2) cannot be an equilibrium point of J_n.

If $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) \geqslant -2h_n'(x_1-x_2)/h_n(x_1-x_2)$, then $u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2) < 2h_n'(x_1-x_2)/h_n(x_1-x_2)$, i.e., $(\partial/\partial x_2)$ log $k_2^n(x_1,x_2) < 0$. Hence (x_1,x_2) cannot be an equilibrium point of J_n for all $n \geqslant 1$. In the case where $u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2) \leqslant u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) \leqslant 0$, we can prove similarly that (x_1,x_2) cannot be an equilibrium point of J_n for all $n \geqslant 1$.

Let us consider the case where $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) < 0$ and $u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2) > 0$. If $u_1'((x_1+x_1)/2)/u_1((x_1+x_2)/2) < -2h_n'(x_1-x_2)/h_n(x_1-x_2)$, then $(\delta/\partial x_1)$ log $k_1^n(x_1,x_2) < 0$, which implies that (x_1,x_2) cannot be an equilibrium point of J_n. If $u_1'((x_1+x_2)/2)/u_2((x_1+x_2)/2) \geqslant -2h_n'(x_1-x_2)/h_n(x_1-x_2)$, then $u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2 < 2h_n'(x_1-x_2)/h_n(x_1-x_2)$, because $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) + u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2) < 0$. Hence we get $(\partial/\partial x_2)$ log $k_2^n(x_1,x_2) < 0$. Hence we have shown that (x_1,x_2) cannot be an equilibrium point of J_n for all $n \geqslant 1$. In the case where $u_1'((x_1+x_2)/2)/u_1((x_1+x_2)/2) > 0$ and $u_2'((x_1+x_2)/2)/u_2((x_1+x_2)/2) < 0$, we can prove similarly that (x_1,x_2) is not be an equilibrium point of J_n for all $n \geqslant 1$. Q.E.D.

Proof of Main Theorem. Initially let us consider the 2nd stage, and let $(q_s,q_b) \in (0,Q]^2$ be arbitrarily chosen and fixed in the 1st stage. We show that $u_1(z) \equiv U(Q-(q_s+q_b)/2, z)$ and $u_2(z) \equiv V((q_s+q_b)/2, M-z)$ on $(0,M]$ satisfy Assumptions (B_1)–(B_4) of Auxiliary Theorem B. $u_1(z)$ and $u_2(z)$ are increasing and decreasing functions by Assumption (A), respectively. Clearly, they are differentiable and concave functions by (B) and (C). It also follows from Lemma 1 that these functions satisfy (B_3). Assumption (B_4) follows from Assumptions (A) and (D). Hence we can apply Auxiliary Theorems A and B to $\{SG(q_s,q_b,g_n)\}$, and it follows that if $(b_s^*,b_b^*) = ((q_s^*,t_s^*),(q_b^*,t_b^*))$ is a perfect equilibrium point, then

$$t_s^*(q_s,q_b) = t_b^*(q_s,q_b) = z^*,$$

where z^* is the number given in (A_3). Of course, this z^* coincides with $m((q_s+q_b)/2)$, which is given in Lemma 2.

Next let us consider the 1st stage. We show that $u_1(q) \equiv u(q) = U(Q-q, m(q))$ and $u_2(q) \equiv v(q) = V(q, M-m(q))$ on $(0,Q]$ satisfy the assumptions of Auxiliary Theorem C. It follows from Lemmas 1 and 2 that $u(q) > 0$ and $v(q) > 0$ for all $q \in (0,Q]$. Assumptions (A_2) and (A_3) follow from Assumption (F) and Lemma 4, respectively. Further (A_1) is true by the following lemma, which will be proved in the Appendix.

LEMMA 5. *$m(q)$ is a differentiable function of q on $(0,Q]$.*

Hence we can apply Auxiliary Theorems A and C to $\{RG(f_n,g_0,t_s^*,t_b^*)\}$, and it follows that if $(b_s^*,b_b^*) = ((q_s^*,t_s^*),(q_b^*,t_b^*))$ is a perfect equilibrium point, then $q_s^* = q_b^* = q_N$, where q_N is given in Lemma 4. Q.E.D.

APPENDIX

Proof of Lemma 1. Let $a = q/Q$. Since $a > 0$, it follows from (C) and (D) that for m given in (D)

$$U(Q-q, am) \geqslant aU(0, m) + (1-a)\, U(w^s) = aU(0, m) > 0,$$
$$V(q, M-am) \geqslant aV(Q, M-m) + (1-a)\, V(w^b) = aV(Q, M-m) > 0.$$

It is clear that $V(q, M) > 0$. Since $U(0, M) > 0$ by (D), we have, by (A), $U(Q-q, M) \geqslant U(0, M) > 0$. For all $m' \in [0, am]$, it holds by (A) that

$$V(q, M-m') \geqslant V(q, M-am) > 0.$$

For all $m' \in [am, M]$, it holds that

$$U(Q-q, m') \geqslant U(Q-q, am) > 0.$$

Q.E.D.

Proof of Lemma 2. For each fixed q, $\log U(Q-q, m) + \log V(q, M-m)$ is defined for some m by Lemma 1. If $m = 0$ or $m = M$, then $U(Q-q, m)\, V(q, M-m) < 0$ by (A) and (D). Hence $\log U + \log V$ is defined on an open interval (a, b) since U and V are concave functions. Since $U(Q-q, a)\, V(q, M-a) = U(Q-q, b)\, V(M-b) = 0$, we have $\lim_{m \to a+0}(\log U + \log V) = \lim_{m \to b-0}(\log U + \log V) = -\infty$. Further since $\log U + \log V$ is a strictly concave function of m, there is a unique $m(q)$ satisfying (1.2). Q.E.D.

Proof of Lemma 3. Let q_1, q_2 and a ($0 < a < 1$, $b = 1-a$ and $q_1 \neq q_2$) be arbitrarily chosen. Then we have, by (C) and definition of $m(q)$,

$$\begin{aligned} aL(q_1) + bL(q_2) &\leqslant \log(aU(Q-q_1, m(q)) + bU(Q-q_2, m(q_2))) \\ &\quad + \log(aV(q_1, M-m(q_1)) + bV(q_2, M-m(q_2))) \\ &< \log U(Q-(aq_1+bq_2), am(q_1)+bm(q_2)) \\ &\quad + \log V(aq_1+bq_2, M-(am(q_1)+bm(q_2))) \\ &\leqslant \log U(Q-(aq_1+bq_2), m(aq_1+bq_2)) \\ &\quad + \log V(aq_1+bq_2, M-m(aq_1+bq_2)) \\ &= L(aq_1+bq_2). \end{aligned}$$

Q.E.D.

Proof of Lemma 4. Since $L(q)$ is a strictly concave function and $\lim_{q \to 0} L(q) = -\infty$, there is a q_N such that $L(q_N) \geqslant L(q)$ for all $q \in (0, Q]$. By (E) we have $q_N < Q$. The uniqueness follows from the strict concavity of $L(q)$.

If there is another allocation $((Q-q^*, m^*), (q^*, M-m^*))$ such that

$q^* \neq q_N$ and $L(q_N) \leqslant \log U(Q-q^*, m^*) + \log V(q^*, M-m^*)$, then it holds by the above that $\log U(Q-q^*, m^*) + \log V(q^*, M-m^*) > L(q^*)$, which is a contradiction. Q.E.D.

Proof of Lemma 5. It follows from Lemma 2 that $m = m(q)$ if and only if

$$\frac{U_2(Q-q, m)}{U(Q-q, m)} = \frac{V_2(q, M-m)}{V(q, M-m)}.$$

Here $U_2 = \partial U/\partial x_2$ and $V_2 = \partial V/\partial x_2$.

Let $f(q, m) = U_2(Q-q, m)/U(Q-q, m) - V_2(q, M-m)/V(q, M-m)$. Then $f(q, m(q)) = 0$. Let us consider $f(q+\Delta q, m(q)+\Delta m) = 0$. Since f has continuous partial derivatives $\partial f/\partial q = f_1$ and $\partial f/\partial m = f_2$ by (B), it holds that

$$\begin{aligned} 0 &= f(q+\Delta q, m(q)+\Delta m) \\ &= f(q+\Delta q, m(q)+\Delta m) - f(q+\Delta q, m(q)) \\ &\quad + f(q+\Delta q, m(q)) - f(q, m(q)) \\ &= f_2(q+\Delta q, m(q)) \cdot \Delta m + o(\Delta m) + f_1(q, m(q)) \cdot \Delta q + o(\Delta q). \end{aligned}$$

Hence we have

$$\frac{\Delta m}{\Delta q}\left(f_2(q+\Delta q, m(q)) + \frac{o(\Delta m)}{\Delta m}\right) + f_1(q, m(q)) + \frac{o(\Delta q)}{\Delta q} = 0.$$

Since $F(m) \equiv \log U(Q-q, m) + \log V(q, M-m)$ is a strictly concave function of m, $d^2F/dm^2 < 0$, i.e., $f_2 < 0$. Hence we have $dm(q)/dq = -f_1/f_2$. Q.E.D.

ACKNOWLEDGMENTS

I wish to express my great thanks to the referee of this journal for helpful suggestions, which made the original proof of the main result much shorter. Of course, any errors that may remain are my responsibility.

REFERENCES

1. J. F. NASH, The bargaining problem, *Econometrica* **18** (1950), 155–162.
2. J. F. NASH, Non-cooperative games, *Ann. of Math.* **54** (1951), 286–296.
3. J. F. NASH, Two-person cooperative games, *Econometrica* **21** (1953), 128–140.
4. R. SELTEN, A simple model of imperfect competition, where 4 are few and 6 are many, *Internat. J. Game Theory* **2** (1973), 141–201.
5. R. SELTEN, Reexamination of the perfectness concept for equilibrium points in extensive games, *Internat. J. Game Theory* **4** (1975), 25–55.

JOURNAL OF ECONOMIC THEORY 24, 328–361 (1981)

Theory of the Firm: Uncertainty and Choice of Experiments

CHIN LIM

Department of Economics, University of Western Ontario,
London, Ontario, Canada

Received July 5, 1978; revised June 11, 1980

1. INTRODUCTION

The basic strategy of a firm comprises the fundamental activities of input purchase, production, and output sales. However, the conduct and organization of these activities differ markedly among firms. Some use price signals, whereas others use quantity signals, for the purpose of transacting in either the input or output market. Furthermore, some firms order their activities whereby input purchase precedes production and output sales; some order their activities in reverse whereby output sales precedes input purchase and production; and others contract their input purchase and output sales simultaneously. In the face of such a large variety of observed behavior, it is both interesting and desirable to understand the environment by which the various strategies can be generated and, more importantly, to explain why one strategy would dominate another.

The analysis begins by considering a firm faces uncertainty in both its output demand and input supply conditions. In the absence of an omniscient auctioneer who freely supplies information on trading prices and the feasible set of trades, the firm in such an environment must necessarily transmit its trade signals in order to discover its trade possibilities. We call such an activity a signalling experiment. A coupling of two signalling experiments, one in each market, together with the final transactions form a complete experiment or, equivalently, a behavioral mode.

In the model, three classes of behavioral modes are generated. The first class characterizes a situation in which the firm first experiments (using either a price or quantity signal) in the input market in order to elicit information on the trading preferences of agents in this market before it makes its input purchase and production decisions. It then experiments (again using either a price or quantity signal) in the output market before sales transactions are conducted. The second class describes a situation where the above sequence of operation is reversed, i.e., the firm first experiments in the

0022-0531/81/030328–34$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

output market and secures a sales contract before it conducts its input market experiment to obtain the required input purchase and production in order to satisfy the prior sales contract. Finally, the third class is one in which the firm simultaneously conducts experiments in the input and output markets before the decisions on input purchase, production, and sales are finalized. Thus, *within* each behavioral class, behavioral modes are differentiated according to whether a price or a quantity signal is used in the input or the output market, and *across* behavioral classes, they are differentiated according to the sequence of experimentation between the input and the output markets.

In the face of the large variety of observed firm behavior, the purpose of this paper is: (1) To demonstrate that the above viewpoint taken at classifying firm behavior provides a useful unifying framework for the understanding of how the large variety of behavior is generated; and (2) to evaluate the relative dominance of modes both *within* and *across* the various behavioral classes. The highlights of our results are: First, the behavioral modes discussed in earlier literature, e.g., Mills [8], Baron [2], Leland [4], Lim [6], and Iwai [3], belong to a small subset of the feasible behavioral modes generated in our model. Also, there is absence of informational advantage á la Blackwell (see Marschak and Miyasawa [7]) between any two modes both *within* and *across* the behavioral classes; hence, behavioral dominance does not trivially depend on the notion of informational superiority. Second, regarding the behavioral dominance *within* each behavioral class: Except for the singular case of decreasing expected marginal cost in the behavioral class where output sales precedes input purchase and production, quantity signals generally dominate price signals. Third, regarding the behavioral dominance between behavioral classes: If uncertainty in the input market is large relative to that in output market, then, subject to some ceteris paribus conditions, the behavioral class where input purchases precede production and output sales is most preferred, whereas if the converse is true, then the most preferred behavioral class is one where output sales precede input purchase and production. Finally, Lim's [6] results on the relative dominance among the behavioral modes considered by Baron [2] and Leland [4], which are but special cases of the three of the twelve behavioral modes considered in this model, are now generalized to include the input market uncertainty.

The rest of the paper is organized as follows. Section 2 describes the information structure facing the firm. Section 3 postulates the behavioral assumptions by which the various behavioral modes are generated. Section 4 describes the nature of the problem in each behavioral mode. Section 5 analyzes the relative dominance between the modes belonging to each behavioral class. In so doing, it addresses the issue of price versus quantity decisions *within* each behavioral class. Section 6 analyzes the relative dominance between the modes *across* the different behavioral classes. The issue here is not only the choice of sequence of experimentation between the input and output markets but also the choice between price and quantity decisions. Finally, a summary and interpretation of the results are presented in Section 7.

2. INFORMATION STRUCTURE OF THE FIRM

Following Leland [4], the firm is assumed to face a stochastic output demand relation which is expressed in the most general form as the implicit function

$$G(q^d, p^d, \eta) = 0, \tag{F(1)}$$

where q^d is quantity demand, p^d is demand price, and $\eta \in [\mathbf{M}, \bar{M}]$ is a random variable, not known ex ante, which has subjective density $d\Psi(\eta)$. Treating F(1) as a mapping from a probability space into a set of downward sloping demand functions, we can place the following restrictions on the implicit function F(1): For any given η, the relation between q^d and p^d is negative; and that larger values of η are associated with larger demand. These restrictions enable us to express F(1) as either

$$q^d = q^d(p^d, \eta), \qquad q_1^d < 0,\ q_2^d > 0, \tag{F(2)}$$

$$p^d = p^d(q^d, \eta), \qquad p_1^d < 0,\ p_2^d > 0. \tag{F(3)}$$

Now define $\eta^0 \equiv E_\eta[\eta]$ and consider two arbitrary values of price, p_α, and quantity, q_α. For sufficiently small uncertainty or random fluctuations of η, expand $q^d(p_\alpha, \eta)$ and $p^d(q_\alpha, \eta)$ around η^0 as

$$q^d(p_\alpha, \eta) \simeq q^d(p_\alpha, \eta^0) + q_2^d(p_\alpha, \eta^0) \cdot (\eta - \eta^0)$$

and

$$p^d(q_\alpha, \eta) \simeq p^d(q_\alpha, \eta^0) + p_2^d(q_\alpha, \eta^0) \cdot (\eta - \eta^0).$$

Taking the expectation of the above, and noting $\eta^0 = E_\eta[\eta]$, $E_\eta[q^d(p_\alpha, \eta)] \simeq q^d(p_\alpha, \eta^0)$ and $E_\eta[p^d(q_\alpha, \eta)] \simeq p^d(q_\alpha, \eta^0)$. These are accurate approximations[1] if the probability distribution is compact and in which variations in η are small (see Samuelson [11]). Now, because $q_1^d < 0$ and

---

[1] The approximations are, of course, exact for the class of demand functions which are linear in the random term η.

$p_1^d < 0$, it follows that for any arbitrary choice of p_α and q_α, and for small uncertainty,

$$q_\alpha \lesseqqgtr E_\eta[q^d(p_\alpha, \eta)] \Leftrightarrow p_\alpha \lesseqqgtr E_\eta[p^d(q_\alpha, \eta)].^2 \qquad \mathrm{F}(4)$$

Analogous to the output demand relation, the stochastic input supply function is also written generally as

$$H(n^s, w^s, \theta) = 0, \qquad \mathrm{F}(5)$$

where n^s is the input supply, w^s is the input supply price or wage, and $\theta \in \Theta$, with interval $[\mathrm{N}, \bar{N}]$, is a random variable, not known ex ante, which has subjective density $d\Phi(\theta)$. The restrictions on F(5) are that n^s and w^s are positively related for any θ, and that large values of θ are associated with larger input supply. These restrictions imply that F(5) can be written as either

$$n^s = n^s(w^s, \theta), \qquad n_1^s > 0,\ n_2^s > 0, \qquad \mathrm{F}(6)$$

or

$$w^s = w^s(n^s, \theta), \qquad w_1^s > 0,\ w_2^s < 0. \qquad \mathrm{F}(7)$$

Analogous to the output demand relation, it can be easily shown that for any arbitrary choice of wage, w_α, and input units, n_α, and for sufficiently small uncertainty,

$$n_\alpha \lesseqqgtr E_\theta[n^s(w_\alpha, \theta)] \Leftrightarrow w_\alpha \gtreqqless E_\theta[w^s(n_\alpha, \theta)]. \qquad \mathrm{F}(8)$$

In addition, it is assumed that θ and η are independently distributed,[3] i.e., the joint distribution

$$\Gamma(\theta, \eta) = \Phi(\theta) \cdot \Psi(\eta). \qquad \mathrm{F}(9)$$

---

[2] This is equivalent to proving $q_\alpha \lesseqqgtr q^d(p_\alpha, \eta^0) \Leftrightarrow p_\alpha \gtreqqless p^d(q_\alpha, \eta^0)$.

Proof. (a) If $q_\alpha \leqslant q^d(p_\alpha, \eta^0)$, then, because $q_1^d < 0$, there must exist a non-negative constant $\varepsilon \geqslant 0$, such that $q_\alpha = q^d(p_\alpha + \varepsilon, \eta^0)$. Since this function is invertible, we can write $p_\alpha + \varepsilon = p^d(q_\alpha, \eta^0) \Rightarrow p_\alpha \leqslant p^d(q_\alpha, \eta^0)$. The converse is true by a similar proof since $p_1^d < 0$.

(b) If $q_\alpha > q^d(p_\alpha, \eta^0)$, then $q_1^d < 0$ implies there exists a positive constant, $\delta > 0$, such that $q_\alpha = q^d(p_\alpha - \delta, \eta^0)$. Since this function is invertible, we can write $p_\alpha - \delta = p^d(q_\alpha, \eta^0) \Rightarrow p_\alpha > p^d(q_\alpha, \eta^0)$. The converse is true by a similar proof since $p_1^d < 0$. Q.E.D.

[3] In general, we have no strong a priori reason to believe whether the output demand and input supply uncertainty should be positively or negatively correlated. The convenient assumption of zero correlation is tantamount to taking the neutral position in the face of equal ignorance. Needless to say, the several results we will obtain rest on the zero correlation assumption and may not be generalized to cover the more difficult cases of non-zero correlation.

Finally, the firm is assumed to know its production function

$$q = f(n), \qquad f'(n) > 0, \tag{F(10)}$$

with certainty, where q is output and n is the input employed.

3. BEHAVIORAL ASSUMPTIONS

The information structure $F(\cdot)$ merely describes the firm's perception of its environment and technology. By itself it does not convey how uncertainty in its output demand and input supply conditions is resolved. Clearly, the behavior of the firm depends critically on the learning mechanism it employs to resolve its uncertainty. In what follows, we postulate the assumptions by which various modes are generated.

A(1). *The revelation of the stochastic variables η and θ is not dependent on time.*

This rules out the degenerate case where the uncertainty problem disappears by simply having the firm waiting till η and θ reveal themselves. In other words, the firm has to always content with a state of imperfect knowledge unless it does the following.

A(2). *The necessary condition for the resolution of uncertainty in the output demand and input supply conditions is that the firm conducts the following signalling experiments: In the output market, it must either signal a supply price, p^s, or quantity supply, q^s; and in the input market, it must either signal a demand wage, w^d, or input demand, n^d.*

In the absence of an omniscient auctioneer who freely supplies information on trading prices and the feasible set of quantity trades, the burden of discovering these rests with the traders themselves. In order to discover its trade possibilities, therefore, the firm must necessarily transmit its trade offer so as to elicit a trade response from its trading agents.

We call the transmission of such information a signal. A signal can either be a price or quantity so that the supply signal variable in the output market is either the supply price, p^s, or quantity supply, q^s; and the demand signal in the input market is either the demand wage, w^d, or the input demand, n^d.

The purpose of a signal is, of course, to elicit a trade response. Such a signal/response relationship is termed a signalling experiment, and the result of which is to establish two pieces of information necessary for trade. Firstly, it establishes the trading price and secondly, it defines the feasible set of quantity trades.

Denoting the symbol → as a signal/response relationship, the signalling experiment in the output market could be either

$$p^s \to q^d(p^s, \eta) \tag{3.1}$$

or

$$q^s \to p^d(q^s, \eta). \tag{3.2}$$

In (3.1), the supply price signal p^s serves also as the trading price set ex ante (η) by the firm and the feasible set of output quantity trades is defined by the response $q^d(p^s, \eta)$, which is the maximum quantity demand registered ex post (η) at the set price. Alternatively, if the quantity supply signal q^s is used as in (3.2), then the trading price is the ex post (η) demand price[4] response $p^d(q^s, \eta)$ and the feasible set of quantity trades is also defined since $p^d(q^s, \eta)$ is the price at which q^s or less is demanded. Similarly, in the input market, the signalling experiment could be either

$$w^d \to n^s(w^d, \theta) \tag{3.3}$$

or

$$n^d \to w^s(n^d, \theta). \tag{3.4}$$

In (3.3), the demand wage signal w^d serves also as the trading price set ex ante (θ) by the firm and the feasible set of input quantity trades is defined by the response $n^s(w^d, \theta)$, which is the maximum available input supply registered ex post (θ) at the set price. Alternatively, if the input quantity demand signal n^d is used as in (3.4), then the trading wage is none other than the ex post (θ) supply wage[5] response $w^s(n^d, \theta)$ and the feasible set of input trades is also defined since $w^s(n^d, \theta)$ indicates the wage at which n^d or less is supplied.

A(3). *Sequential signalling experimentation is ruled out because of prohibitively high experimental costs.*

Now suppose sequential signalling experimentation in each market is ruled out because of prohibitively high experimental costs so that the firm in each market is restricted to a single signalling experiment. Then it is necessary that each experiment constitutes an actual bilateral trade offer on which basis actual transactions are conducted. An experiment is therefore used here only as a heuristic device to illuminate the nature of the firm's real decision rule under uncertainty.

---

[4] While the notion of a quantity demand reponse to a supply price signal is straightforward, that of a demand price response to a quantity supply signal poses a little difficulty. With regard to the latter, we may have to require the existence of some marketing or an intermediate trading agency that quotes a demand price to the firm's quantity offer. Such intermediate agencies might not exist for all markets. Wherever they do not exist, there is still a useful and interesting implication to be drawn (as we will later observe) by treating the quantity signalling behavior as a conceptual exercise.

[5] The argument in the preceding footnote applies here but in the context of the input market.

A(4). *The nature of feasible contracts ex post to each signalling experiment.*

A transaction is characterized basically by a contract specifying a transacting price and quantity. Let q and n be the feasible quantity transactions in the output and input markets respectively. Then the feasible contracts ex post to the signalling experiments in (3.1) to (3.4) are, respectively,

$$(p^s, q), \quad \text{where } q \leqslant q^d(p^s, \eta), \tag{3.1$'$}$$

$$(p^d(q^s, \eta), q), \quad \text{where } q \leqslant q^s, \tag{3.2$'$}$$

$$(w^d, n), \quad \text{where } n \leqslant n^s(w^d, \theta), \tag{3.3$'$}$$

$$(w^s(n^d, \theta), n), \quad \text{where } n \leqslant n^d. \tag{3.4$'$}$$

The feasible contracts corresponding to the price signalling experiments are immediately apparent. In experiment (3.1): $p^s \to q^d(p^s, \eta)$, the signal p^s also serves as the transacting price, and the response $q^d(p^s, \eta)$ defines the maximum available demand registered ex post (η) at price p^s; hence contract (3.1)′. In experiment (3.3): $w^d \to n^s(w^d, \theta)$, the signal w^d represents also the transacting wage, and the response $n^s(w^d, \theta)$ defines the maximum input supply registered ex post (θ) at wage w^d; hence contract (3.3)′.

When quantity signalling experiments are used, however, the nature of their corresponding feasible contracts are as follows. In experiment (3.2): $q^s \to p^d(q^s, \eta)$, the signal q^s is the ex ante (η) offer of quantity supply. But since the ex post (η) demand price response $p^d(q^s, \eta)$ registers the transacting price at which q^s *or less* is demanded, the firm in the ex post (η) contract (3.2)′ need not commit itself to sell according to its signal q^s, but may entertain a transaction $q < q^s$ if, for instance, $p^d(q^s, \eta)$ turns out to be too low. Similarly, in experiment (3.4): $n^d \to w^s(n^d, \theta)$, the signal n^d is the amount of input it offers to purchase ex ante (θ). But since the ex post (θ) supply wage response $w^s(n^d, \theta)$ represents the transacting wage at which n^d *or less* is supplied, the firm in the ex post (θ) contract (3.4)′ need not commit itself to purchase at an input level according to its signal n^d, but may entertain a transaction $n < n^d$ if, for instance, $w^s(n^d, \theta)$ turns out to be too high.

In all the above contracts, (3.1)′ to (3.4)′, the transacting price in each contract is established either by the firm as a price signal (if a price signalling experiment is used), or by the response price (if a quantity signalling experiment is used). Assumption A(3) rules out the possibility of trading at prices other than the single price observation revealed within the single experiment.

A(5). *The information structure* $F(\cdot)$ *does not vary with the type of signalling experiment.*

The information set $F(\cdot)$ is the only one available to the firm. There is no other. This forms the basis by which we can legitimately analyze the relative dominance of various experiments.

A(6). *An experiment is chosen to maximize expected profit*[6] *of a firm that produces a perishable good.*

4. THE FORMAL PROBLEM IN EACH EXPERIMENT

Since the function of a firm is to buy (in the input market) and sell (in the output market), a complete experiment consists of a coupling of two signalling experiments, one conducted in each market. Denoting I (either n^d or w^d) and O (either q^s or p^s) respectively as an input and output market signal, three classes of experiments can be logically generated and characterized by the following ordered pairs: (I, O), (O, I) and $\binom{I}{O}$.

(a) The class of (I, O) behavior, which contains four possible experiments in (n^d, q^s), (w^d, q^s), (n^d, p^s), and (w^d, p^s), is one where signalling experiment O is conducted *after* experiment I is completed.

(b) The class of (O, I) behavior, which contains experiments (q^s, n^d), (q^s, w^d), (p^s, n^d), and (p^s, w^d), is one where signalling experiment I is conducted *after* experiment O is completed.

(c) The class of $\binom{I}{O}$ behavior, which contains experiments $\binom{n^d}{q^s}$, $\binom{w^d}{q^s}$, $\binom{n^d}{p^s}$, and $\binom{w^d}{p^s}$, is one where both signalling experiments I and O are conducted *simultaneously*.

Defining:

ex ante (θ) or (η) variable $\equiv$ a choice variable made *prior* to the revelation of θ or η,

and

---

[6] The analysis quickly becomes complicated for the case where firms are risk-averse. Obviously, our results rest on the risk neutrality assumption and extensions to risk-averse firms await further analysis.

ex post (θ) or (η) variable $\equiv$ a choice variable made *after* the revelation of θ or η,

the formal problem in each class of experiments is described as follows.

4.1. *The (I, O) Problem*

The (I, O) behavior can be described as the class of "buy-then-sell" activities. Faced with uncertainty in both the input supply and output demand conditions, the firm begins its operation by signalling I in the input market. Upon receiving a trade response from this market, it then makes its actual input purchase, n, which uniquely determines the output level, q, via the production function. It subsequently signals O in the output market in order to transact sales. Now since I is ex ante (θ) and (η) and n, q and O are ex post (θ) but ex ante (η), the formal problem, which is to maximize expected profit, is characterized by[7]

$$\underset{I}{\text{Maximize}}\, E_\theta \underset{n,q,O}{\text{Maximum}}\, E_\eta[\pi(I, \theta, n, q, O, \eta)]. \tag{4.1}$$

The description of each experiment and its formal problem is presented below and also summarized in Table I.

(a) (n^d, q^s) *behavior.*

In this behavior, note that $q = f(n)$, and $q^s = q$, so that the maximizing problem can be reduced to

$$\underset{n^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^d}{\text{Maximum}}\, R_q(n) - w^s(n^d, \theta) \cdot n, \tag{4.2}$$

where

$$R_q(n) \equiv E_\eta[p^d(f(n), \eta)] \cdot f(n) \tag{4.3}$$

is the expected revenue product function when quantity, q^s, is used as the output signal. The solution to the above problem can be found by backward induction from the following two-stage procedure.

---

[7] The arguments in the profit function $\pi(\cdot)$ for each subsequent behavioral class are ordered in sequence to express the sequential steps by which each choice variable is made. For example, in $\pi(I, \theta, n, q, O, \eta)$, I is chosen ex ante (θ) and (η) whereas n, q, O are chosen ex post (θ) but ex ante (η). Also note that the profit functions generally differ among behavioral classes. However, notational complexity makes it desirable to denote $\pi(\cdot)$ as the profit function which implicitly differs among behavioral classes according to the arguments contained within.

TABLE I

The (I, O) Problem

	Behavior	Input market signalling experiment	Employment and production control	Output market signalling experiment	Ex post profit: $\pi(I, \theta, n, q, O, \eta)$	Formal problem: $\underset{I}{\text{Maximize}}\, E_\theta \underset{n,q,O}{\text{Maximum}}\, E_\eta[I, \theta, n, q, O, \eta]$
(a)	(n^d, q^s)	$n^d \to w^s(n^d, \theta)$	$n \leqslant n^d$ $\Rightarrow q = f(n)$	$(q^s = q)$ $q^s \to p^d(q^s, \eta)$	$p^d(q^s, \eta) \cdot q$ $- w^s(n^d, \theta) \cdot n$	$\underset{n^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^d, q, q^s}{\text{Maximum}}\, E_\eta[p^d(q^s, \eta) \cdot q - w^s(n^d, \theta) \cdot n]$ yields solution $a_*^\theta \equiv (n^d_{a^*}, n^\theta_{a^*})$
(b)	(w^d, q^s)	$w^d \to n^s(w^d, \theta)$	$n \leqslant n^s(w^d, \theta)$ $\Rightarrow q = f(n)$	$(q^s = q)$ $q^s \to p^d(q^s, \eta)$	$p^d(q^s, \eta) \cdot q - w^d \cdot n$	$\underset{w^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^s(w^d, \theta), q, q^s}{\text{Maximum}}\, E_\eta[p^d(q^s, \eta) \cdot q - w^d \cdot n]$ yields solution $b_*^\theta \equiv (w^d_{b^*}, n^\theta_{b^*})$
(c)	(n^d, p^s)	$n^d \to w^s(n^d, \theta)$	$n \leqslant n^d$ $\Rightarrow q = f(n)$	$p^s \to q^d(p^s, \eta)$	$p^s \cdot \text{Min}(q, q^d(p^s, \eta))$ $- w^s(n^d, \theta) \cdot n$	$\underset{n^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^d, q, p^s}{\text{Maximum}}\, E_\eta[p^s \cdot \text{Min}(q, q^d(p^s, \eta)) - w^s(n^d, \theta) \cdot n]$ yields solution $c_*^\theta \equiv (n^d_{c^*}, n^\theta_{c^*})$
(d)	(w^d, p^s)	$w^d \to n^s(w^d, \theta)$	$n \leqslant n^s(w^d, \theta)$ $\Rightarrow q = f(n)$	$p^s \to q^d(p^s, \eta)$	$p^s \cdot \text{Min}(q, q^d(p^s, \eta))$ $- w^d \cdot n$	$\underset{w^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^s(w^d, \theta), q, p^s}{\text{Maximum}}\, E_\eta[p^s \cdot \text{Min}(q, q^d(p^s, \eta)) - w^d \cdot n]$ yields solution $d_*^\theta \equiv (w^d_{d^*}, n^\theta_{d^*})$

(i) Given signal n^d, and hence the response $w^s(n^d, \theta)$, the first stage is to choose n so as to

$$\underset{n \leqslant n^d}{\text{maximize}}\, R_q(n) - w^s(n^d, \theta) \cdot n. \tag{4.4}$$

It is assumed that

$$R_q'(n) > 0 \qquad \text{and} \qquad R_q''(n) \leqslant 0 \qquad \text{for positive } n. \tag{4.5}$$

The solution to the above can be generally expressed as[8]

$$n_a^\theta \equiv n_a(n^d, \theta) \leqslant n^d \qquad \text{for any } \theta. \tag{4.6}$$

(ii) The next stage is to determine optimal signal n^d. Substituting $n_a(n^d, \theta)$ for n in (4.4) and then taking its expectation with respect to θ, optimal n^d is found as follows.

$$\underset{n^d}{\text{Maximize}}\, E_\theta [R_q(n_a(n^d, \theta)) - w^s(n^d, \theta) \cdot n_a(n^d, \theta)]. \tag{4.7}$$

Let $n_{a^*}^d$ be the solution to (4.7). Then the optimal input purchase is

$$n_{a^*}^\theta \equiv n_a(n_{a^*}^d, \theta) \leqslant n_{a^*}^d \qquad \text{for any } \theta. \tag{4.8}$$

The entire solution can be characterized as

$$a_*^\theta \equiv (n_{a^*}^d, n_{a^*}^\theta) \tag{4.9}$$

and the maximum expected profit

$$E[\pi_a(a_*^\theta)] \equiv E_\theta [R_q(n_{a^*}^\theta) - w^s(n_{a^*}^d, \theta) \cdot n_{a^*}^\theta]. \tag{4.10}$$

(b) (w^d, q^s) *behavior.*

The maximizing problem in this behavior, which can be reduced to

$$\underset{w^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^s(w^d, \theta)}{\text{Maximum}}\, R_q(n) - w^d \cdot n, \tag{4.11}$$

is similar to the preceding problem, (4.2), except for the input signal, and the input purchase constraint. Following the two-stage procedure of solving for w^d and n, the entire solution can thus be generally written as

$$b_*^\theta \equiv (w_{b^*}^d, n_{b^*}^\theta), \tag{4.12}$$

---

[8] From now on, the notation for the solution follows the alphabetical order of indexing of each experiment. For example, the letter "a" is used to denote the solution in the (n^d, q^s) behavior which is indexed under subheading "(a)."

where $w_{b^*}^d$ is the optimal demand wage signal and the optimal input purchase

$$n_{b^*}^{\theta} \equiv n_b(w_{b^*}^d, \theta) \leqslant n^s(w_{b^*}^d, \theta) \qquad \text{for any } \theta. \tag{4.13}$$

The maximum expected profit

$$E[\pi_b(b_*^{\theta})] \equiv E_\theta[R_q(n_{b^*}^{\theta}) - w_{b^*}^d \cdot n_{b^*}^{\theta}]. \tag{4.14}$$

(c) (n^d, p^s) *behavior.*

Using $q \equiv f(n)$, the maximizing problem in this behavior is reduced to

$$\begin{aligned} &\underset{n^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^d, p^s}{\text{Maximum}}\, E_\eta[p^s \cdot \text{Min}(f(n), q^d(p^s, \eta)) - w^s(n^d, \theta) \cdot n] \\ &= \underset{n^d}{\text{Maximize}}\, E_\theta \underset{n \leqslant n^d, p^s}{\text{Maximum}}\, R(n, p^s) - w^s(n^d, \theta) \cdot n, \end{aligned} \tag{4.15}$$

where $R(n, p^s)$ is the expected sales revenue which can be expressed as either

$$R(n, p^s) = p^s f(n) + p^s \cdot V(n, p^s) \tag{4.16}$$

or

$$R(n, p^s) = p^s E_\eta[q^d(p^s, \eta)] + p^s \cdot S(n, p^s), \tag{4.17}$$

where

$$V(n, p^s) \equiv \int_{M}^{\eta^*} [q^d(p^s, \eta) - f(n)]\, d\Psi(\eta) \leqslant 0 \tag{4.18}$$

and

$$S(n, p^s) \equiv \int_{\eta^*}^{\bar{M}} [f(n) - q^d(p^s, \eta)]\, d\Psi(\eta) \leqslant 0 \tag{4.19}$$

and where η^* satisfies $q^d(p^s, \eta^*) = f(n)$. Note that $V(n, p^s)$ and $S(n, p^s)$ are non-positive for any combination of n and p^s. $p^s V(n, p^s)$ in (4.16) is the expected loss in revenue due to unsold inventory whenever demand falls short of production.[9] $p^s \cdot S(n, p^s)$ in (4.17) is the expected loss in revenue due to foregone potential sales whenever demand exceeds production. Thus (4.16) or (4.17) implies that the expected sales revenue, $R(n, p^s)$, cannot exceed the value of production $p^s \cdot f(n)$, or the expected value of demand,

---

[9] In this single period model in which the good is assumed perishable, all unsold inventory yields a zero net value and is therefore considered a total loss. In a multi-period model, however, and whenever the good is durable, unsold inventory usually has a positive net value equal to its expected value minus the storage cost, and hence while the loss due to unsold inventory still exists, it is less severe than in the case considered in this model.

$p^s \cdot E_\eta[q^d(p^s, \eta)]$, respectively. The solution to **problem (4.15)**, using the *two-stage procedure*, proceeds as follows.

(i) *For any given input signal, n^d, and the response, $w^s(n^d, \theta)$, the first stage is to determine n and p^s. First, to solve for p^s given n,*

$$\underset{p^s}{\text{Maximize}}\, R(n, p^s) - w^s(n^d, \theta) \cdot n. \tag{4.20}$$

Since p^s does not appear in the cost term, $w^s(n^d, \theta) \cdot n$, the solution to (4.20) can be expressed generally as the expected sales revenue maximizing price, $p^s(n)$, for any given n. Substituting $p^s(n)$ for p^s in (4.20), n is then chosen to

$$\underset{n \leq n^d}{\text{Maximize}}\, R_p(n) - w^s(n^d, \theta) \cdot n, \tag{4.21}$$

where $R_p(n) \equiv R(n, p^s(n))$. It is assumed that

$$R'_p(n) > 0 \qquad \text{and} \qquad R''_p(n) \leq 0 \qquad \text{for positive } n. \tag{4.22}$$

The input purchase solution can then be generally expressed as

$$n_c^\theta \equiv n_c(n^d, \theta) \leq n^d \qquad \text{for any } \theta. \tag{4.23}$$

(ii) The next stage is to determine n^d. Substituting $n_c(n^d, \theta)$ for n in (4.21) and then taking its expectation with respect to θ, optimal n^d is found as follows.

$$\underset{n^d}{\text{Maximize}}\, E_\theta[R_p(n_c(n^d, \theta)) - w^s(n^d, \theta) \cdot n_c(n^d, \theta)]. \tag{4.24}$$

Let $n_{c^*}^d$ be the optimal solution to (4.24). Then the optimal input purchase,

$$n_{c^*}^\theta \equiv n_c(n_{c^*}^d, \theta) \leq n_{c^*}^d \qquad \text{for any } \theta \tag{4.25}$$

and the entire solution can be written as

$$c_*^\theta \equiv (n_{c^*}^d, n_{c^*}^\theta). \tag{4.26}$$

The maximum expected profit is then

$$E[\pi_c(c_*^\theta)] = E_\theta[R_p(n_{c^*}^\theta) - w^s(n_{c^*}^d, \theta) \cdot n_{c^*}^\theta]. \tag{4.27}$$

(d) (w^d, p^s) *behavior.*

The maximizing problem in this behavior

$$\underset{w^d}{\text{Maximize}}\, E_\theta \underset{n \leq n^s(w^d, \theta), p^s}{\text{Maximum}}\, R(n, p^s) - w^d \cdot n \tag{4.28}$$

is similar to that of (n^d, p^s) behavior in (4.15) except for the input signal and the input purchase constraint. Thus, the solution to (4.28) is easily inferred from that of (4.15) and can be generally expressed as

$$d_*^\theta \equiv (w_{d\cdot}^d, n_{d\cdot}^\theta),$$

where $w_{d\cdot}^d$ is the optimal demand wage *signal and the* optimal input purchase,

$$n_{d\cdot}^\theta \equiv n_d(w_{d\cdot}^d, \theta) \leqslant n^s(w_{d\cdot}^d, \theta) \qquad \text{for any } \theta. \tag{4.29}$$

The maximum expected profit

$$E[\pi_d(d_*^\theta)] \equiv E_\theta[R_p(n_{d\cdot}^\theta) - w_{d\cdot}^d \cdot n_{d\cdot}^\theta]. \tag{4.30}$$

The price adjustment model of Iwai [3] is identically the (w^d, p^s) behavior considered above. Also, both the quantity-cum-price setting behavior (originated by Mills [8]) and the quantity-setting behavior considered by Baron [2], Leland [4], and Lim [6], although they do not explicitly consider the uncertainty in the input market, are analogous to the above (n^d, p^s) and (n^d, q^s) behavior, respectively.[10]

Furthermore, it is important to note that the expected revenue product function, $R_q(\cdot)$, which obtains when q^s is used as the output signal differs from $R_p(\cdot)$, which issues from the experiment when p^s is used as the signal. The latter contains the element of potential loss in unsold inventory which is absent in the former.

4.2. *The (O, I) Problem*

The (O, I) behavior can be described as the class of "sell-then-buy" activities. Assuming that a forward market exists, the firm begins its operation by announcing its sales signal 0. Upon receiving a trade response, it then makes its actual sales transaction, q, which it seeks to satisfy by purchasing, via the inverse of the production function, the required level of input, n. The input market signal I is therefore constrained by the prior sales contract only to n^d $(=n)$ since the alternate signal w^d might give rise to inconsistency whenever the response $n^s(w^d, \theta) < n$. Note that in the (O, I) problem, O is ex ante (η) and (θ) whereas q, n, and I are ex post (η) but ex ante (θ) so that the formal problem is characterized as

$$\underset{O}{\text{Maximize}}\ E_\eta \underset{q,n,I}{\text{Maximum}}\ E_\theta[\pi(O, \eta, q, n, I, \theta)]. \tag{4.31}$$

---

[10] It should be noted that the firm is assumed to be risk averse in Baron [2] and Leland [4] but risk neutral in Iwai [3], Mills [8], and Lim [6].

TABLE II

The (O, I) Problem

Behavior	Output market signalling experiment	Supply control and implied employment control	Input market signalling experiment	Ex post profit: $\pi(O, \eta, q, n, I, \theta)$	Formal problem: $\text{Maximize}_{O}\ E_\eta \ \text{Maximum}_{q,n,I}\ E_\theta[\pi(O, \eta, q, n, I, \theta)]$
(a) (q^s, n^d)	$q^s \to p^d(q^s, \eta)$	$q \leqslant q^s$ $\Rightarrow n = f^{-1}(q) \equiv h(q)$	$(n^d = n)$ $n^d \to w^s(n^d, \theta)$	$p^d(q^s, \eta) \cdot q$ $- w^s(h(q), \theta) \cdot h(q)$	$\text{Maximize}_{q^s}\ E_\eta \ \text{Maximum}_{q \leqslant q^s, n, n^d}\ E_\theta[p^d(q^s, \eta) \cdot q$ $- w^s(h(q), \theta) \cdot h(q)]$ yields solution $a^\eta_* \equiv (q^s_{a^*}, q^\eta_{a^*})$
(b) (q^s, w^d)	$q^s \to p^d(q^s, \eta)$	$q \leqslant q^s$ $\Rightarrow n = f^{-1}(q) \equiv h(q)$	$w^d \to n^s(w^d, \theta)$	possible inconsistency whenever $n^s(w^d, \theta) < n$	
(c) (p^s, n^d)	$p^s \to q^d(p^s, \eta)$	$q \leqslant q^d(p^s, \eta)$ $\Rightarrow n = f^{-1}(q) \equiv h(q)$	$(n^d = n)$ $n^d \to w^s(n^d, \theta)$	$p^s \cdot q$ $- w^s(h(q), \theta) \cdot h(q)$	$\text{Maximize}_{p^s}\ E_\eta \ \text{Maximum}_{q \leqslant q^d(p^s, \eta), n, n^d}\ E_\theta[p^s \cdot q$ $- w^s(h(q), \theta) \cdot h(q)]$ yields solution $c^\eta_* \equiv (p^s_{c^*}, q^\eta_{c^*})$
(d) (p^s, w^d)	$p^s \to q^d(p^s, \eta)$	$q \leqslant q^d(p^s, \eta)$ $\Rightarrow n = f^{-1}(q) \equiv h(q)$	$w^d \to n^s(w^d, \theta)$	possible inconsistency whenever $n^s(w^d, \theta) < n$	

The description of each experiment and its formal problem is presented below and also summarized in Table II. These experiments can be further simplified because once sales, q, is transacted, this uniquely determines the required input signal and employment as $n^d = n = f^{-1}(q) \equiv h(q)$ so that the prior sales commitment is satisfied.

(a) *(q^s, n^d) behavior.*

The maximizing problem in this behavior can be reduced to

$$\underset{q^s}{\text{Maximize}}\; E_\eta \underset{q \leqslant q^s}{\text{Maximum}}\; p^d(q^s, \eta) \cdot q - C(q), \tag{4.32}$$

where

$$C(q) \equiv E_\theta[w^s(h(q), \theta)] \cdot h(q) \tag{4.33}$$

is the expected cost function defined on q. Note that $C'(q) > 0$. The solution to (4.32) is found by the following two-stage procedure:

(i) Given any sales signal, q^s, and hence the response, $p^d(q^s, \eta)$, q is chosen to

$$\underset{q \leqslant q^s}{\text{Maximize}}\; p^d(q^s, \eta) \cdot q - C(q). \tag{4.34}$$

The solution to (4.34) can be expressed generally as

$$q_a^\eta \equiv q_a(q^s, \eta) \leqslant q^s \qquad \text{for any } \eta. \tag{4.35}$$

(ii) The next stage is to determine optimal signal, q^s. Substituting $q_a(q^s, \eta)$ for q in (4.34) and then taking its expectation with respect to η, q^s is chosen to

$$\underset{q^s}{\text{Maximize}}\; E_\eta[p^d(q^s, \eta) \cdot q_a(q^s, \eta) - C(q_a(q^s, \eta))]. \tag{4.36}$$

Let $q_{a^*}^s$ be the solution to (4.36), then the optimal sales transaction,

$$q_{a^*}^\eta \equiv q_a(q_{a^*}^s, \eta) \leqslant q_{a^*}^s \qquad \text{for any } \eta. \tag{4.37}$$

The entire solution can be generally written as

$$a_*^\eta \equiv (q_{a^*}^s, q_{a^*}^\eta) \tag{4.38}$$

and the maximum expected profit

$$E[{}_a\pi(a_*^\eta)] \equiv E_\eta[p^d(q_{a^*}^s, \eta) \cdot q_{a^*}^\eta - C(q_{a^*}^\eta)]. \tag{4.39}$$

(c) *(p^s, n^d) behavior.*

The maximizing problem in this behavior

$$\underset{p^s}{\text{Maximize}}\ E_\eta \underset{q \leqslant q^d(p^s,\eta)}{\text{Maximum}}\ p^s \cdot q - C(q) \tag{4.40}$$

is similar to (4.32) except for the sales signal and the sales constraint. Thus, the solution to (4.40) is easily inferred from that of (4.32) and can be generally expressed as

$$c^\eta_* \equiv (p^s_{c^*}, q^\eta_{c^*}), \tag{4.41}$$

where $p^s_{c^*}$ is the optimal supply price signal and the optimal sales,

$$q^\eta_{c^*} \equiv q_c(p^s_{c^*}, \eta) \leqslant q^d(p^s_{c^*}, \eta) \qquad \text{for any } \eta. \tag{4.42}$$

The maximum expected profit

$$E[{}_c\pi(c^\eta_*)] \equiv E_\eta[p^s_{c^*} \cdot q^\eta_{c^*} - C(q^\eta_{c^*})]. \tag{4.43}$$

The price-setting behavior considered by Baron [2], Leland [4], and Lim [6] is analogous to the (p^s, n^d) behavior considered here.

4.3. *The* $\binom{I}{O}$ *Problem*

The $\binom{I}{O}$ behavior can be described as the class of "buy-and-sell" activities. Both the input signal I and output signal O are announced simultaneously. Following the trade responses from both markets, the firm then conducts its output sales, q, and the required input purchase, n. Now, both I and O are ex ante (θ) and (η) whereas n and q are ex post (θ) and (η) so that the formal problem is to

$$\underset{I,O}{\text{Maximize}}\ E_{\theta,\eta} \underset{n,q}{\text{Maximum}}\ \pi(I, O, \theta, \eta, n, q). \tag{4.44}$$

The description of each experiment and its formal problem is summarized in Table III. The solution to each experiment again follows the two-stage procedure utilized previously. As an example, the solution to $\binom{n^d}{q^s}$ is described below.

(A) $\binom{n^d}{q^s}$ *behavior*

The solution to this maximizing problem

$$\underset{n^d,q^s}{\text{Maximize}}\ E_{\theta,\eta} \underset{n \leqslant \text{Min}[n^d,h(q^s)]}{\text{Maximum}}\ p^d(q^s, \eta) \cdot f(n) - w^s(n^d, \theta) \cdot n \tag{4.45}$$

can be generally written as

$$A^{\theta,\eta}_* = (n^d_{A^*}, q^s_{A^*} = f(n^d_{A^*}), n^{\theta,\eta}_{A^*}), \tag{4.46}$$

TABLE III

The $\binom{I}{O}$ Problem

Behavior	Simultaneous input and output signalling experiments	Controls	Remarks	Formal problem: $\underset{I,O}{\text{Maximize}}\, E_{\theta,\eta}\, \underset{n,q}{\text{Maximum}}\, \pi(I, O, \theta, \eta, n, q)$
(A) $\binom{n^d}{q^s}$	$n^d \to w^s(n^d, \theta)$ $q^s \to p^d(q^s, \eta)$	$n \leqslant n^d$ $q \leqslant q^s$	Since both control constraints must apply simultaneously, we must have $n \leqslant \text{Min}[n^d, h(q^s)]$	$\underset{n^d, q^s}{\text{Maximize}}\, E_{\theta,\eta} \underset{n \leqslant \text{Min}[n^d, h(q^s)]}{\text{Maximum}} p^d(q^s, \eta) \cdot f(n) - w^s(n^d, \theta) \cdot n$ yields solution $A_*^{\theta,\eta} \equiv (n^d_{A\cdot}, q^s_{A\cdot} = f(n^d_{A\cdot}), n^{\theta,\eta}_{A\cdot})$
(B) $\binom{w^d}{q^s}$	$w^d \to n^s(w^d, \theta)$ $q^s \to p^d(q^s, \eta)$	$n \leqslant n^s(w^d, \theta)$ $q \leqslant q^s$	Since both control constraints must apply simultaneously, we must have $n \leqslant \text{Min}[n^s(w^d, \theta), h(q^s)]$	$\underset{w^d, q^s}{\text{Maximize}}\, E_{\theta,\eta} \underset{n \leqslant \text{Min}[n^s(w^d,\theta), h(q^s)]}{\text{Maximum}} p^d(q^s, \eta) \cdot f(n) - w^d \cdot n$ yields solution $B_*^{\theta,\eta} \equiv (w^d_{B\cdot}, q^s_{B\cdot}, n^{\theta,\eta}_{B\cdot})$
(C) $\binom{n^d}{p^s}$	$n^d \to w^s(n^d, \theta)$ $p^s \to q^d(p^s, \eta)$	$n \leqslant n^d$ $q \leqslant q^d(p^s, \eta)$	Since both control constraints must apply simultaneously, we must have $n \leqslant \text{Min}[n^d, h(q^d(p^s, \eta))]$	$\underset{n^d, p^s}{\text{Maximize}}\, E_{\theta,\eta} \underset{n \leqslant \text{Min}[n^d, h(q^d(p^s,\eta))]}{\text{Maximum}} p^s \cdot f(n) - w^s(n^d, \theta) \cdot n$ yields solution $C_*^{\theta,\eta} \equiv (n^d_{C\cdot}, p^s_{C\cdot}, n^{\theta,\eta}_{C\cdot})$
(D) $\binom{w^d}{p^s}$	$w^d \to n^s(w^d, \theta)$ $p^s \to (p^s, \eta)$	$n \leqslant n^s(w^d, \theta)$ $q \leqslant q^d(p^s, \eta)$	Since both control constraints must apply simultaneously, we must have $n \leqslant \text{Min}[n^s(w^d, \theta), h(q^d(p^s, \eta))]$	$\underset{w^d, p^s}{\text{Maximize}}\, E_{\theta,\eta} \underset{n \leqslant \text{Min}[n^s(w^d,\theta), h(q^d(p^s,\eta))]}{\text{Maximum}} p^s \cdot f(n) - w^d \cdot n$ yields solution $D_*^{\theta,\eta} \equiv (w^d_{D}, p^s_{D\cdot}, n^{\theta,\eta}_{D\cdot})$

where $n_A^d\cdot$ and $q_A^s\cdot$ are the optimal signal variables and

$$n_A^{\theta;\eta} \equiv n_A(n_A^d\cdot, q_A^s\cdot, \theta, \eta) \leqslant \mathrm{Min}[n_A^d\cdot, h(q_A^s\cdot)] \text{ for any } \theta, \eta, \tag{4.47}$$

is the optimal input employment. The maximum expected profit

$$E[\pi_A(A_*^{\theta,\eta})] \equiv E_{\theta,\eta}[p^d(q_A^s\cdot, \eta)\cdot f(n_A^{\theta;\eta}) - w^s(n_A^d\cdot, \theta)\cdot n_A^{\theta;\eta}]. \tag{4.48}$$

The solutions to (B) $\binom{w^d}{q^s}$, (C) $\binom{n^d}{p^s}$, and (D) $\binom{w^d}{p^s}$ behavior are analogous to the above and which differ only in the signals and constraints. Using the appropriate notation, these solutions are adequately summarized in Table III.

From the informational standpoint, there appears to be no apparent advantages between experiments *within* and *across* the various behavioral classes. Consider the (I, O) behavior with its ordered choice vector $(I, \theta, n, q, O, \eta)$; the (O, I) behavior with its ordered choice vector $(O, \eta, q, n, I, \theta)$; and the $\binom{I}{O}$ behavior with its choice vector $(I, O, \theta, \eta, n, q)$. *Within* each behavioral class, it is apparent there is no informational advantage between any two experiments since they share an equal number of ex ante and ex post decision variables. Also, any informational advantage *between* behavioral classes does not seem to exist. This is clear between the (I, O) and (O, I) behavior since their choice vectors reveal each has to make an equal number of decisions ex ante and ex post. Between the (I, O) (it suffices to discuss either (I, O) or (O, I)) behavior and the $\binom{I}{O}$ behavior, the apparent nonexistence of informational advantage between them arises as follows. On the one hand, (I, O) seems to be informationally superior over $\binom{I}{O}$ because in the latter behavior, two decisions, I and O, are made ex ante (θ) and ex ante (η), whereas in the former behavior, only decision I is ex ante (θ) and ex ante (η) but decision O is ex post (θ). On the other hand, (I, O) is informationally inferior to $\binom{I}{O}$ because the decisions, n and q, in the latter behavior can be made ex post (θ) and ex post (η) whereas in the former behavior, they are made ex post (θ) but ex ante (η). The absence of informational advantage between experiments *within* and *across* the behavioral classes suggests, therefore, that behavioral dominance cannot, in general, be based trivially on Blackwell's criterion of informational superiority. Instead it relies on other conditions as will be demonstrated in the following sections.

5. RELATIVE DOMINANCE OF WITHIN CLASS EXPERIMENTS

The experiments within each behavioral class are differentiated according to whether quantity or price is used as a market signal. The analysis of choice of within class experiments is therefore a study of comparative advantage between quantity and price signalling experiments.

5.1. *Relative Dominance of Experiments Within the Class of (I, O) Behavior*

For sufficiently small uncertainty such that F(4) and F(8) hold, the results on the relative dominance of experiments within the (I, O) class of behavior are as follows.

PROPOSITION 5.1. *(n^d, q^s) behavior is preferred to (n^d, p^s) behavior.*

Proof. (n^d, p^s): By definition, $c_*^\theta = (n_{c\cdot}^d, n_{c\cdot}^\theta)$ is optimal. The ex post (θ) expected profit in (4.27), after using (4.16) and (4.17), can be expressed as either

$$\pi_c(c_*^\theta) = p^s(n_{c\cdot}^\theta) \cdot f(n_{c\cdot}^\theta) + p^s(n_{c\cdot}^\theta) \cdot V(n_{c\cdot}^\theta, p^s(n_{c\cdot}^\theta)) - w^s(n_{c\cdot}^d, \theta) \cdot n_{c\cdot}^\theta \tag{5.1}$$

or

$$\begin{aligned}\pi_c(c_*^\theta) = {} & p^s(n_{c\cdot}^\theta) \cdot E_\eta[q^d(p^s(n_{c\cdot}^\theta), \eta)] + p^s(n_{c\cdot}^\theta) \cdot S(n_{c\cdot}^\theta, p^s(n_{c\cdot}^\theta)) \\ & - w^s(n_{c\cdot}^d, \theta) \cdot n_{c\cdot}^\theta.\end{aligned} \tag{5.2}$$

Now, depending on the nature of demand, cost, and the probability distribution functions, the relationship between optimal supply price, $p^s(n_{c\cdot}^\theta)$, and output, $f(n_{c\cdot}^\theta)$, can only take either of the following forms, i.e., (i) $E_\eta[q^d(p^s(n_{c\cdot}^\theta), \eta)] \geqslant f(n_{c\cdot}^\theta)$ or (ii) $E_\eta[q^d(p^s(n_{c\cdot}^\theta), \eta)] \leqslant f(n_{c\cdot}^\theta)$, which states that for any state of nature θ, the firm could either face expected excess demand or excess supply respectively. Hence we define $\hat{\theta} \in \Theta$ such that

$$E_\eta[q^d(p^s(n_{c\cdot}^{\hat{\theta}}), \eta)] \geqslant f(n_{c\cdot}^{\hat{\theta}}) \tag{5.3}$$

and $\bar{\theta} \in \Theta$ such that

$$E_\eta[q^d(p^s(n_{c\cdot}^{\bar{\theta}}), \eta)] \leqslant f(n_{c\cdot}^{\bar{\theta}}). \tag{5.4}$$

(n^d, p^s) vs (n^d, q^s) behavior: (i) *If random outcome $\hat{\theta}$*, then in (n^d, p^s) behavior, the optimal solution is, by definition, $c_*^{\hat{\theta}} = (n_{c\cdot}^d, n_{c\cdot}^{\hat{\theta}})$, and maximum ex post $(\hat{\theta})$ expected profit, using (5.1), is

$$\begin{aligned}\pi_c(c_*^{\hat{\theta}}) = {} & p^s(n_{c\cdot}^{\hat{\theta}}) \cdot f(n_{c\cdot}^{\hat{\theta}}) + p^s(n_{c\cdot}^{\hat{\theta}}) \cdot V(n_{c\cdot}^{\hat{\theta}}, p^s(n_{c\cdot}^{\hat{\theta}})) \\ & - w^s(n_{c\cdot}^d, \hat{\theta}) \cdot n_{c\cdot}^{\hat{\theta}}.\end{aligned} \tag{5.5}$$

In (n^d, q^s) behavior, let a feasible solution $a^{\hat{\theta}} = (n_{c\cdot}^d, n_{c\cdot}^{\hat{\theta}})$ be chosen. Substituting $a^{\hat{\theta}}$ for a_*^θ in (4.10), the ex post $(\hat{\theta})$ expected profit, after using (4.3), is

$$\pi_a(a^{\hat{\theta}}) = E_\eta[p^d(f(n_{c\cdot}^{\hat{\theta}}), \eta)] \cdot f(n_{c\cdot}^{\hat{\theta}}) - w^s(n_{c\cdot}^d, \hat{\theta}) \cdot n_{c\cdot}^{\hat{\theta}}. \tag{5.6}$$

Subtracting (5.5) from (5.6) yields

$$\pi_a(a^{\hat\theta}) - \pi_c(c_*^{\hat\theta}) = \{E_\eta[p^d(f(n_{c^*}^{\hat\theta}), \eta)] - p^s(n_{c^*}^{\hat\theta})\} \cdot f(n_{c^*}^{\hat\theta}) - p^s(n_{c^*}^{\hat\theta}) \cdot V(n_{c^*}^{\hat\theta}, p^s(n_{c^*}^{\hat\theta})) \geqslant 0. \tag{5.7}$$

The above inequality holds because, firstly, $V(\cdot) \leqslant 0$, and secondly (5.3) and F(4) imply $E_\eta[p^d(f(n_{c^*}^{\hat\theta}), \eta)] \geqslant p^s(n_{c^*}^{\hat\theta})$.

(ii) *If random outcome* $\bar\theta$, then in (n^d, p^s) behavior, the optimal solution is $c_*^{\bar\theta} = (n_{c^*}^d, n_{c^*}^{\bar\theta})$ and the maximum ex post $(\bar\theta)$ expected profit, using (5.2) is

$$\pi_c(c_*^{\bar\theta}) = p^s(n_{c^*}^{\bar\theta}) \cdot E_\eta[q^d(p^s(n_{c^*}^{\bar\theta}), \eta)] + p^s(n_{c^*}^{\bar\theta}) \cdot S(n_{c^*}^{\bar\theta}, p^s(n_{c^*}^{\bar\theta})) - w^s(n_{c^*}^d, \bar\theta) \cdot n_{c^*}^{\bar\theta}. \tag{5.8}$$

In (n^d, q^s) behavior, let a feasible solution, $a^{\bar\theta} = (n_{c^*}^d, n_a^{\bar\theta})$, be chosen such that $n_a^{\bar\theta}$ satisfies

$$f(n_a^{\bar\theta}) = E_\eta[q^d(p^s(n_{c^*}^{\bar\theta}), \eta)]. \tag{5.9}$$

The solution $a^{\bar\theta}$ is feasible on account of (5.9), (5.4), (4.25), and (4.6). Furthermore (5.9) and F(4) imply

$$E_\eta[p^d(f(n_a^{\bar\theta}), \eta] = p^s(n_{c^*}^{\bar\theta}). \tag{5.10}$$

With this choise of $a^{\bar\theta}$, the ex post $(\bar\theta)$ expected profit in (n^d, q^s) behavior is

$$\pi_a(a^{\bar\theta}) = E_\eta[p^d(f(n_a^{\bar\theta}), \eta)] \cdot f(n_a^{\bar\theta}) - w^s(n_{c^*}^d, \bar\theta) \cdot n_a^{\bar\theta}. \tag{5.11}$$

Subtracting (5.8) from (5.11), and using (5.9) and (5.10), yields

$$\pi_a(a^{\bar\theta}) - \pi_c(c_*^{\bar\theta}) = -p^s(n_{c^*}^{\bar\theta}) \cdot S(n_{c^*}^{\bar\theta}, p^s(n_{c^*}^{\bar\theta})) + w^s(n_{c^*}^d, \bar\theta) \cdot [n_{c^*}^{\bar\theta} - n_a^{\bar\theta}] \geqslant 0. \tag{5.12}$$

The above inequality holds because, firstly, $S(\cdot) \leqslant 0$ and secondly, (5.9) and (5.4) imply $f(n_{c^*}^{\bar\theta}) \geqslant f(n_a^{\bar\theta}) \Rightarrow n_{c^*}^{\bar\theta} \geqslant n_a^{\bar\theta}$.

The inequalities (5.7) and (5.12) imply that for any state of nature θ, $\pi_a(a^\theta) \geqslant \pi_c(c_*^\theta) \Rightarrow E[\pi_a(a^\theta)] \geqslant E[\pi_c(c_*^\theta)] \Rightarrow E[\pi_a(a_*^\theta)] \geqslant E[\pi_c(c_*^\theta)]$ since a_*^θ is by definition optimal in (n^d, q^s) behavior. Q.E.D.

By analogously applying the above proof, the following result is easily established.

PROPOSITION 5.2. *(w^d, q^s) behavior is preferred to (w^d, p^s) behavior.*

Both Propositions 5.1 and 5.2 establish that the dominance of a quantity signal over a price signal in the output market. That the same holds true for input market signals is shown by the following results.

PROPOSITION 5.3. *(n^d, q^s) behavior is preferred to (w^d, q^s) behavior.*

Proof. *(w^d, q^s) behavior*: By definition, $b_*^\theta = (w_{b\cdot}^d, n_{b\cdot}^\theta)$ is optimal. The maximum expected profit, (4.14), is rewritten as

$$E[\pi_b(b_*^\theta)] = E_\theta[R_q(n_{b\cdot}^\theta)] - w_{b\cdot}^d \cdot E_\theta[n_{b\cdot}^\theta]. \tag{5.13}$$

Note that because of inequality in (4.13),

$$E_\theta[n_{b\cdot}^\theta] \leqslant E_\theta[n^s(w_{b\cdot}^d, \theta)]. \tag{5.14}$$

(w^d, q^s) vs. (n^d, q^s) behavior: In (n^d, q^s) behavior, let a feasible solution, $a = (n_a^d, n_a^d)$, be chosen such that

$$n_a^d = E_\theta[n_{b\cdot}^\theta]. \tag{5.15}$$

Substituting solution a for a_*^θ in (4.10), the expected profit is

$$E[\pi_a(a)] = E_\theta[R_q(n_a^d) - w^s(n_a^d, \theta) \cdot n_a^d]. \tag{5.16}$$

Subtracting (5.13) from (5.16), and then using (5.15), yields

$$\begin{aligned} E[\pi_a(a)] - E[\pi_b(b_*^\theta)] = {} & \{R_q(E_\theta[n_{b\cdot}^\theta]) - E_\theta[R_q(n_{b\cdot}^\theta)]\} \\ & + \{w_{b\cdot}^d - E_\theta[w^s(E_\theta[n_{b\cdot}^\theta], \theta)]\} \cdot E_\theta[n_{b\cdot}^\theta] \geqslant 0. \end{aligned} \tag{5.17}$$

The above inequality holds because firstly, $R_q(\cdot)$ is concave and therefore by Jensen's inequality, $R_q(E_\theta[n_{b\cdot}^\theta]) \geqslant E_\theta[R_q(n_{b\cdot}^\theta)]$, and secondly, (5.14) and F(8) imply $w_{b\cdot}^d \geqslant E_\theta[w^s(E_\theta[n_{b\cdot}^\theta], \theta)]$. Q.E.D.

By analogously applying the above proof, but this time using the assumption that $R_p(\cdot)$ is concave, the following result is easily established.

PROPOSITION 5.4. *(n^d, p^s) behavior is preferred to (w^d, p^s) behavior.*

Finally, both Propositions 5.1 and 5.4 or Propositions 5.2 and 5.3 imply

PROPOSITION 5.5. *(n^d, q^s) behavior is preferred to (w^d, p^s) behavior.*

Quantity signals in the (I, O) behavioral class, therefore, generally dominate price signals. This occurs for two separate reasons as can be inferred from the proofs. Firstly, the concavity of the expected revenue product functions defined in the input quantity space, $R_q(\cdot)$ and $R_p(\cdot)$, imply, by Jensen's inequality, that expected revenue (and hence expected profit) is *lower, ceteris paribus, the greater the uncertainty in quantities of input supply. Thus, the input quantity signal dominates the input price signal since the latter, but not the former, elicits uncertain quantity responses of input* supply. Secondly, in the (I, O) behavior, the output is already produced before sales are conducted. Thus, the use of quantity as an output sales signal (which results in the expected revenue product function defined by $R_q(\cdot)$) dominates the use of an output price signal (in which case $R_p(\cdot)$ is the expected revenue product function) simply because $R_p(\cdot)$, but not $R_q(\cdot)$, contains the element of possible loss due to unsold inventories whenever demand falls short of production.

5.2. *Relative Dominance of Experiments Within the Class of (O, I) Behavior*

In the (O, I) behavioral class, the feasible experiments as shown in Table II are reduced to the (q^s, n^d) and (p^s, n^d) experiments. It is shown below that the relative dominance between them depends critically on whether the expected marginal cost is an increasing or decreasing function of output.

PROPOSITION 5.6. *For increasing (decreasing) expected marginal cost firms, (q^s, n^d) behavior is preferred (dispreferred) to (p^s, n^d) behavior.*

Proof. (i) *The case of increasing expected marginal cost*: In (p^s, n^d) behavior, $c^\eta_* = (p^s_{c^*}, q^\eta_{c^*})$ is, by definition, optimal and the maximum expected profit is as in (4.43). In (q^s, n^d) behavior, let a feasible solution, $a = (q^s_a, q^s_a)$, be chosen such that

$$q^s_a = E_\eta[q^\eta_{c^*}]. \tag{5.18}$$

Substituting solution a for a^η_* in (4.39),

$$E[{}_a\pi(a)] = E_\eta[p^d(q^s_a, \eta)] \cdot q^s_a - C(q^s_a). \tag{5.19}$$

Subtracting (4.43) from (5.19), and using (5.18), yields

$$\begin{aligned} E[{}_a\pi(a)] - E[{}_c\pi(c^\eta_*)] = {} & \{E_\eta[p^d(E_\eta[q^\eta_{c^*}], \eta)] - p^s_{c^*}\} \cdot E_\eta[q^\eta_{c^*}] \\ & + \{E_\eta[C(q^\eta_{c^*})] - C(E_\eta[q^\eta_{c^*}])\} > 0. \end{aligned}$$

The above inequality holds because firstly, the inequality in (4.42) implies $E_\eta[q^\eta_{c^*}] \leqslant E_\eta[q^d(p^s_{c^*}, \eta)]$ which together with F(4) imply $E_\eta[p^d(E_\eta[q^\eta_{c^*}], \eta)] \geqslant p^s_{c^*}$; and secondly, $C'' > 0$ and by Jensen's inequality, $E_\eta[C(q^\eta_{c^*})] > C(E_\eta[q^\eta_{c^*}])$.

(ii) *The case of decreasing expected marginal cost*: In this case, $C'' < 0$ and hence the optimal solution in (q^s, n^d) behavior, $a^\eta_* \equiv (q^s_{a^*}, q^\eta_{a^*} = q^s_{a^*})$ since it is clear that optimal sales $q^\eta_{a^*} = q^s_{a^*}$ because the firm would exploit its decreasing expected marginal cost condition by making a maximum permissible sales contract. Using $a^\eta_* = (q^s_{a^*}, q^s_{a^*})$, the maximum expected profit in (4.39) is modified to

$$E[{}_a\pi(a^\eta_*)] \equiv E_\eta[p^d(q^s_{a^*}, \eta)] \cdot q^s_{a^*} - C(q^s_{a^*}). \tag{5.20}$$

In (p^s, n^d) behavior, let a feasible solution, $c^\eta = (p^s_c, q^d(p^s_c, \eta))$, be chosen such that

$$p^s_c = E_\eta[p^d(q^s_{a^*}, \eta)], \tag{5.21}$$

which from F(4) implies

$$E_\eta[q^d(p^s_c, \eta)] = q^s_{a^*}. \tag{5.22}$$

Substituting c^η for c^η_* in (4.43),

$$E[{}_c\pi(c^\eta)] = p^s_c \cdot E_\eta[q^d(p^s_c, \eta)] - E_\eta[C(q^d(p^s_c, \eta))]. \tag{5.23}$$

Subtracting (5.20) from (5.23), and using (5.21) and (5.22), yields

$$E[{}_c\pi(c^\eta)] - E[{}_a\pi(a^\eta_*)] = C(E_\eta[q^d(p^s_c, \eta)]) - E_\eta[C(q^d(p^s_c, \eta))] > 0$$

owing to Jensen's inequality because $C'' < 0$. Q.E.D.

The above results are now explained. If the expected cost function is convex, then increasing uncertainty in output demand, by Jensen's inequality, would, ceteris paribus, increase the expected cost and hence decrease the expected profit. Conversely, increasing output demand undertainty would, ceteris paribus, increase the expected profit level if the expected cost function is concave. Thus, in the case of increasing expected marginal cost (or convex expected cost function), a quantity signal dominates a price signal since the latter, but not the former, elicits uncertain responses in quantity demand. The converse reasoning easily establishes the dominance of the price over the quantity signal if the expected marginal cost is decreasing.

5.3. *Relative Dominance of Experiments Within the Class of* $\binom{I}{O}$ *Behavior*

The relative dominance between experiments in the $\binom{I}{O}$ behavioral class is easily inferred in light of earlier results. Observe that in the (I, O) behavioral class, the input quantity dominates the input price signal when the expected revenue product functions are concave. Also note that in the (O, I) behavioral class, the output quantity dominates the output price signal whenever the expected cost function is convex. Now since the $\binom{I}{O}$ behavioral class is a sort of a hybrid between the (I, O) and (O, I) behavioral class, it is expected that the conditions governing the dominance of experiments within the latter two behavioral classes should carry over to determine the dominance of experiments within the $\binom{I}{O}$ behavior. In particular, if the expected revenue product function is concave and the expected cost function is convex, then the use of quantity signals in either the input or output market should dominate the use of price signals. This is stated as

PROPOSITION 5.7. *If the expected marginal revenue product is diminishing and the expected marginal cost is increasing, then* $\binom{n^d}{q^s}$ *behavior is preferred to either* $\binom{w^d}{q^s}$ *or* $\binom{n^d}{p^s}$ *behavior which in turn is preferred to* $\binom{w^d}{p^s}$ *behavior.*

6. RELATIVE DOMINANCE OF ACROSS CLASS EXPERIMENTS

In this section we turn to examine the relative dominance between experiments belonging to different classes of behavior. In particular, we consider the two following distinct issues.

Firstly, since the comparison of experiments belonging to different behavioral classes involves experiments which may differ in their input and output signal combinations, it is desirable to isolate from the effects of price versus quantity decisions which have been extensively studied in the preceding section and focus on experiments with a similar input and output signal combination. For instance, the experiments, $(n^d, q^s) \in (I, O)$, $(q^s, n^d) \in (O, I)$, and $\binom{n^d}{q^s} \in \binom{I}{O}$, have a similar input and output combination but they belong to different behavioral classes. Abstracting from the effects of price versus quantity decisions, the conditions governing the preferences of one behavioral class over another is best illustrated by considering the following two diametrically extreme situations where uncertainty exists in either one or the other market only. Defining O_u as a situation where uncertainty exists in the output demand but not in the input supply condition and I_u as a situation in reverse, then it is readily verified[11] that

PROPOSITION 6.1. *Under* $I_u(O_u)$, $(n^d, q^s) \in (I, O)$ *behavior is preferred (dispreferred) to* $\binom{n^d}{q^s} \in \binom{I}{O}$ *behavior, which in turn is preferred (dispreferred) to* $(q^s, n^d) \in (O, I)$ *behavior.*

This result is obvious from the following illustration. Consider the ordered choice vectors: $(I, \theta, n, q, O, \eta) \in (I, O)$ behavior; $(I, O, \theta, \eta, n, q) \in \binom{I}{O}$ behavior; and $(O, \eta, q, n, I, \theta) \in (O, I)$ behavior. Without output market uncertainty, the variable η can be deleted as irrelevant and the above decision vectors reduce to $(I, \theta, n, q, O) \in (I, O)$ behavior; $(I, O, \theta, n, q) \in \binom{I}{O}$ behavior; and $(O, q, n, I, \theta) \in (O, I)$ behavior. Now, comparing the number of decision variables that have to be made ex ante (θ) and ex post (θ) between the behavioral classes clearly reveals that (I, O) behavior is more informative (or ex post (θ) more flexible) than $\binom{I}{O}$ behavior which in turn is more informative than (O, I) behavior. Thus, informational advantage between behavioral classes (which previously has been shown to be absent in the general case where uncertainty exists in both markets) can now be established for the polar case where uncertainty exists in only one or the other market. In the above example where uncertainty is present only in the input market, therefore, Blackwell's criterion for behavioral dominance based on the principle of greater informativeness clearly establishes $(n^d, q^s) \in (I, O)$ to be preferred to $\binom{n^d}{q^s} \in \binom{I}{O}$ which in turn is preferred to $(q^s, n^d) \in (O, I)$ behavior. In the opposite polar case where uncertainty exists only in output market, a similar reasoning would verify that the dominance is in the reverse direction. In other words, if uncertainty exists only in the input supply but not the output demand condition, then it is optimal to use an (n^d, q^s) strategy to experiment first in the input market and resolve any further uncertainty in this market by making an input purchase contract. This strategy leaves the firm with no uncertainty at all when it makes its output sales. In contrast, the (q^s, n^d) strategy, which involves experimenting first in the output market, is inferior since the absence of uncertainty in the output market makes any experiment redundant and not only so, the firm has to bear the full brunt of the input market uncertainty when it makes its input purchase to satisfy a prior sales contract. If, however, uncertainty exists in the output, but not the input market, the converse of the above argument easily establishes the optimality of the (q^s, n^d) strategy. From this, we conjecture that if uncertainty were to exist in both markets, then the (n^d, q^s) srategy is most preferred if uncertainty in the input market is large relative to that in the output market whereas the (q^s, n^d) strategy is most preferred if the converse is true.

---

[11] The proof of Proposition 6.1 is straightforward and is contained in Lim [5].

It is important to emphasize that Proposition 6.1 may not be generalized to the group of behavior: $(n^d, p^s) \in (I, O)$, $\binom{n^d}{p^s} \in \binom{I}{O}$, and $(p^s, n^d) \in (O, I)$. For instance, in spite of the greater informativeness of $(p^s, n^d) \in (O, I)$ over $(n^d, p^s) \in (I, O)$ behavior under O_u (uncertainty existing in the output market only), it is generally not true that the former would always be preferred to the latter behavior. Proposition 6.4 below, which is trivially applicable under O_u, shows that (n^d, p^s) may, under some circumstances, dominate (p^s, n^d) behavior. The above anomaly arises because the profit *functions* $\pi_a(\cdot)$, $\pi_A(\cdot)$ and ${}_a\pi(\cdot)$ corresponding to the group of behavior: (n^d, q^s), $\binom{n^d}{q^s}$ and (q^s, n^d), are identical whereas the profit functions $\pi_c(\cdot)$, $\pi_C(\cdot)$ and ${}_c\pi(\cdot)$ corresponding the group of behavior: (n^d, p^s), $\binom{n^d}{p^s}$ and (p^s, n^d), in general, differ. In particular, recall that implicit in the $\pi_c(\cdot)$ function is an element of loss due to unsold inventory which is absent in either $\pi_C(\cdot)$ or ${}_C\pi(\cdot)$. Hence Blackwell's criterion may not be generally applicable to establish dominance among the latter group of behavior.[12]

Secondly, the two issues of price versus quantity decisions and the decisions on which market to experiment first are now considered simultaneously. Essentially, it involves the examination of behavioral dominance between experiments which not only belong to different behavioral classes but which also differ in their input and output signal combinations. We shall focus on three experiments, $(n^d, q^s) \in (I, O)$, $(p^s, n^d) \in (O, I)$, and $(n^d, p^s) \in (I, O)$, which are respectively analogous to the quantity-setting, price-setting, and quantity-cum-price setting behavior discussed in earlier literature by Baron [2], Leland [4], and Lim [6]. Why this literature has restricted the feasible behavioral set to the above three modes is a direct consequence of the fact the input market and its uncertainty have been ignored. The analogy between the above three experiments and the ones considered in the literature becomes apparent when we consider the following illustration. Recall the $(n^d, q^s) \in (I, O)$ behavior with its choice vector $(n^d, \theta, n, q, q^s, \eta)$; the $(p^s, n^d) \in (O, I)$ behavior with its choice vector $(p^s, \eta, q, n, n^d, \theta)$; and the $(n^d, p^s) \in (I, O)$ behavior with its choice vector $(n^d, \theta, n, q, p^s, \eta)$. Now if we ignore the input market and its uncertainty, the input market variables n^d, n, and θ become irrelevant and can be deleted. This reduces the choice vectors to: (q, q^s, η) for the $(n^d, q^s) \in (I, O)$ behavior; (p^s, η, q) for the $(p^s, n^d) \in (O, I)$ behavior; and (q, p^s, η) for the $(n^d, p^s) \in (I, O)$ behavior. Now based on the ex ante (η) choice variables in the above choice vectors, the behavioral modes (n^d, q^s), (p^s, n^d) and (n^d, p^s) can be appropriately termed, as in the literature, the quantity-setting, price-setting, and quantity-cum-price setting behavior. Furthermore, examining the number of ex post (η) choice variables reveal the $(p^s, n^d) \in (O, I)$ (price-setting) behavior, which contains one ex post (η) choice variable in q, to be more informative or ex post (η) more flexible than either the $(n^d, q^s) \in (I, O)$ (quantity-setting) behavior or the $(n^d, p^s) \in (I, O)$ (quantity-cum-price setting) behavior, which contains no ex post (η) choice variable.[13]

---

[12] Let γ be a random variable and x, y, z be decision variables affecting the pay-off functions $J(\cdot)$ and $J'(\cdot)$. By Blackwell's criterion, $\text{Max}_x E_\gamma[\text{Max}_{y,z} J(\cdot)] \geq \text{Max}_{x,y} E_\gamma[\text{Max}_z J'(\cdot)]$ provided $J(\cdot) = J'(\cdot)$. Otherwise, if $J(\cdot) \neq J'(\cdot)$ then the above inequality does not necessarily hold.

[13] This contrasts with Leland's [4] characterization which puts the quantity-setting and the price-setting behavior to be both ex post (η) more flexible than the quantity-cum-price setting behavior. We are not exactly clear what implicit assumptions are involved in his model. In our model, however, if uncertainty is absent in the input market, then the quantity-setting behavior (with choice vector (q, q^s, η)) and the quantity-cum-price setting behavior (with choice vector (q, p^s, η)) both belong to the (I, O) class where production, q, is committed *before* sales is conducted. In either case, whether q^s or p^s is used as a sales signal, the firm ex post (η) has no choice but to absorb whatever demand condition that is realized. They are

Nevertheless, as has been before argued, greater informativeness in this case is insufficient to establish behavioral dominance since the profit functions differ among these three behavioral modes. This claim is apparent from the set of results below (see especially Propositions 6.2 and 6.4). These results are trivially true under O_u and, therefore, are but generalizations of Lim's [6] results to include input market uncertainty.

PROPOSITION 6.2. *For increasing expected marginal cost,* $(n^d, q^s) \in (I, O)$ *is preferred to* $(p^s, n^d) \in (O, I)$ *behavior.*

Proof. In (p^s, n^d) behavior, $c^{\eta}_{*} \equiv (p^s_{c^*}, q^{\eta}_{c^*})$ is, by definition, optimal and the maximum expected profit is as in (4.43), and rewritten as

$$E[{}_c\pi(c^{\eta}_{*})] \equiv p^s_{c^*} . E_{\eta}[q^{\eta}_{c^*}] - E_{\eta}[C(q^{\eta}_{c^*})]. \tag{6.1}$$

In (n^d, q^s) behavior, let a feasible solution, $a = (n^d_a, n^d_a)$, be chosen such that $f(n^d_a) = E_{\eta}[q^{\eta}_{c^*}]$. Now, since $E_{\eta}[q^{\eta}_{c^*}] \leqslant E_{\eta}[q^d(p^s_{c^*}, \eta)]$ and because of F(4), we have

$$E_{\eta}[p^d(f(n^d_a), \eta)] \geqslant p^s_{c^*}. \tag{6.2}$$

With the choice of solution a, the expected profit in (n^d, q^s) behavior,

$$E[\pi_a(a)] = E_{\eta}[p^d(f(n^d_a), \eta)]\, f(n^d_a) - E_{\theta}[w^s(n^d_a, \theta)] \cdot n^d_a. \tag{6.3}$$

Subtracting (6.1) from (6.3), and noting that $n^d_a = h(E_{\eta}[q^{\eta}_{c^*}])$ and the definition of $C(\cdot)$ in (4.33),

$$\begin{aligned} E[\pi_a(a)] - E[{}_c\pi(c^{\eta}_{*})] = {} & \{E_{\eta}[p^d(f(n^d_a), \eta)] - p^s_{c^*}\}\, f(n^d_a) \\ & + \{E_{\eta}[C(q^{\eta}_{c^*})] - C(E_{\eta}[q^{\eta}_{c^*}])\} > 0. \end{aligned}$$

The above inequality is due to (6.2) and Jensen's inequality since $C(\cdot)$ is strictly convex. Q.E.D.

PROPOSITION 6.3. *For increasing expected marginal cost,* $(n^d, q^s) \in (I, O)$ *is preferred to* $(n^d, p^s) \in (I, O)$ *behavior.*

The above follows from Proposition 5.1 which holds irrespective of whether expected marginal cost is increasing or decreasing.

PROPOSITION 6.4. *For increasing expected marginal cost,* $(p^s, n^d) \in (O, I)$ *is not always preferred to* $(n^d, p^s) \in (I, O)$ *behavior.*

---

both, therefore, not as ex post (η) flexible as the price-setting behavior (with choice vector (p^s, η, q)) which belongs to the (O, I) class because, here, production, q, proceeds only *after* the demand condition is realized which gives this behavior an extra degree of ex post (η) flexibility.

Proof. In (p^s, n^d) behavior, $c_*^\eta \equiv (p_{c\cdot}^s, q_{c\cdot}^\eta)$ is, by definition, optimal and the maximum expected profit is (6.1). In (n^d, p^s) behavior, $c_*^\theta = (n_{c\cdot}^d, n_{c\cdot}^\theta, p^s(n_{c\cdot}^\theta))$ is, by definition, optimal. Nevertheless, let an arbitrary feasible solution, $c = (n_c^d, n_c^d, p_{c\cdot}^s)$, be chosen such that

$$f(n_c^d) = E_\eta[q_{c\cdot}^\eta] \Rightarrow n_c^d = h(E_\eta[q_{c\cdot}^\eta]). \tag{6.4}$$

Substituting c for c_*^θ in (4.27) and using (4.16), the expected profit from the choice of solution c,

$$E[\pi_c(c)] = p_{c\cdot}^s\{f(n_c^d) + V(n_c^d, p_{c\cdot}^s)\} - E_\theta[w^s(n_c^d, \theta)]\, n_c^d. \tag{6.5}$$

Subtracting (6.1) from (6.5) yields, after using (6.4) and noting the definition of $C(\cdot)$ in (4.33),

$$E[\pi_c(c)] - E[{}_c\pi(c_*^\eta)] = p_{c\cdot}^s V(n_c^d, p_{c\cdot}^s) + \{E_\eta[C(q_{c\cdot}^\eta)] - C(E_\eta[q_{c\cdot}^\eta])\}, \tag{6.6}$$

which cannot be unambiguously signed since in the R.H.S. of (6.6), the first term is non-positive and the second term is positive by Jensen's inequality since $C(\cdot)$ is strictly convex. Since $V(\cdot)$ strictly depends on the nature of the demand condition which is totally unrelated to the cost condition, there exists some demand and cost conditions whereby (6.6) is non-negative. In an example considered by Lim [6] involving a quadratic cost function and a linear demand function, it was shown that for any given variance in demand, (6.6) tends to be positive the larger the slope of the marginal cost function.

Q.E.D.

The ambiguity in the dominance relationship between (n^d, p^s) and (p^s, n^d) experiments is due to the following. First, in Proposition 6.3, (n^d, q^s) dominates (n^d, p^s) behavior for the reason provided earlier which is essentially the disadvantage of the latter in having to absorb a possible loss due to unsold invertories. Second, in Proposition 6.2, (n^d, q^s) dominates (p^s, n^d) behavior because the latter, but not the former, strategy elicits random quantity responses in output demand which puts it at a disadvatage in light of the convexity of the expected cost function. Thus, in Proposition 6.4, the dominance between (n^d, p^s) and (p^s, n^d) behavior is ambiguous depending on the sizes of the above two off-setting disadvantages. Proposition 6.4 trivially applies under O_u (uncertainty existing only in the output market). Thus under O_u, (p^s, n^d) does not generally dominate (n^d, p^s); hence it confirms our previous claim that Proposition 6.1 does not generalize to the group of behavior: (n^d, p^s), $({}^{n^d}_{p^s})$, and (p^s, n^d).

For decreasing expected marginal cost, due to the difficulty of establishing general results, we shall nevertheless consider the cases where the input purchase contract for the (n^d, q^s) behavior is bound by the input signal, i.e., $n = n^d$, and the sales contract for the (p^s, n^d) behavior is bound by the response to the output signal, i.e., $q = q^d(p^s, \eta)$. For these plausible cases, we have

PROPOSITION 6.5. *For decreasing expected marginal cost,* $(p^s, n^d) \in (O, I)$, *is preferred to* $(n^d, q^s) \in (I, O)$ *which in turn is preferred to* $(n^d, p^s) \in (I, O)$ *behavior.*

Proof. In (n^d, q^s) behavior, $a_*^\theta \equiv (n_{a\cdot}^d, n_{a\cdot}^\theta)$, is by definition, optimal. If the input purchase is constrained to be equal to the input signal, then $n_{a\cdot}^\theta = n_{a\cdot}^d$, and the maximum expected profit in (4.10) can be rewritten as

$$E[\pi_a(a_*^\theta)] = E_\eta[p^d(f(n_{a\cdot}^d), \eta)]\, f(n_{a\cdot}^d) - E_\theta[w^s(n_{a\cdot}^d, \theta)]\, n_{a\cdot}^d. \tag{6.7}$$

In (p^s, n^d) behavior, let a feasible solution, $c^n \equiv (p_c^s, q^d(p_c^s, \eta))$, be chosen such that

$$p_c^s = E_\eta[p^d(f(n_{a\cdot}^d), \eta)], \tag{6.8}$$

which from F(4) implies

$$E_\eta[q^d(p_c^s, \eta)] = f(n_{a\cdot}^d) \Rightarrow n_{a\cdot}^d = h(E_\eta[q^d(p_c^s, \eta)]). \tag{6.9}$$

The expected profit in (p^s, n^d) behavior from the choice of solution c^n,

$$E[{}_c\pi(c^n)] = p_c^s E_\eta[q^d(p_c^s, \eta)] - E_\eta[C(q^d(p_c^s, \eta))]. \tag{6.10}$$

Subtracting (6.7) from (6.10) yields, after noting (6.8), (6.9), and the definition of $C(\cdot)$ in (4.33)

$$E[{}_c\pi(c^n)] - E[\pi_a(a_*^\theta)] = C(E_\eta[q^d(p_c^s, \eta)]) - E_\eta[C(q^d(p_c^s, \eta))] > 0$$

because of Jensen's inequality since $C(\cdot)$ is strictly concave. Thus for decreasing expected marginal cost, (p^s, n^d) is preferred to (n^d, q^s) behavior which in turn is preferred to (n^d, p^s) behavior by Proposition 5.1. Q.E.D.

The preceding result is easily explained. Firstly, $(p^s, n^d) \in (O, I)$ is preferred to $(n^d, q^s) \in (I, O)$ for the simple reason that a concave expected cost function, by Jensen's inequality, favors a strategy where production is ex post (η) variable (a characteristic of (p^s, n^d) behavior) over one in which production cannot be varied ex post (η) (a characteristic of (n^d, q^s)). Secondly, $(n^d, q^s) \in (I, O)$ is in turn preferred to $(n^d, p^s) \in (I, O)$ for the reason provided earlier, which is that the latter, but not the former, strategy incurs an element of loss due to unsold inventories.

7. Concluding Remarks

In summary, we have shown how a firm facing both input supply and output demand uncertainty may adopt various strategies differentiated according to the timing of experimentation, using price or quantity signals, between the input and output markets. A behavior or experiment, therefore, could differ from another depending on (a) whether price of quantity signals are used in the experiment in each market, and on (b) whether it belongs to the (I, O) strategy (where the input signalling precedes the output signalling experiment), or the (O, I) strategy (where the output signalling precedes the input signalling experiment) or the $\binom{I}{O}$ strategy (where both the input and output signalling experiments are conducted simultaneously). All these experiments seem to be equally informative á la Blackwell and hence the dominance relationship among them cannot appeal simply to the criterion of information superiority. Instead the relative dominance of experiments *within* and *across* the behavioral classes is due to the following issues.

On the *first issue* of price versus quantity decisions, it is certainly not apparent whether or not price should dominate quantity signals. The reason is that in a price signalling experiment, the firm has to absorb uncertainty in quantity responses, but so must it absorb uncertainty of price responses in a quantity signalling experiment. We have, however, been able to obtain some striking results which are the following.

On the one hand, we found quantity to dominate price signals whenever the following conditions occur: diminishing expected marginal revenue product (or concave expected revenue product functions) in the (I, O) strategy; increasing expected marginal cost (or convex expected cost functions) in the (O, I) strategy; and diminishing expected marginal revenue product together with increasing expected marginal cost in the $\binom{I}{O}$ strategy. To explain these results, observe that the concavity of the expected revenue product function (defined on the input quantity space) and the convexity of the expected cost function (defined on the output quantity space) imply, by Jensen's inequality, that expected profit would be smaller, ceteris paribus, the greater the uncertainty in quantities of input supply and of output demand, respectively. For this reason and also the fact that prices, but not quantity signals, elicit uncertain quantity responses in input supply or output demand, it follows that in the (I, O) strategy, the input quantity dominates the input price signals; in the (O, I) strategy, the output quantity dominates the output price signal; and in the $\binom{I}{O}$ strategy, both input and output quantity signal dominate the price signals. What remains to be explained is the dominanc of the quantity over the price signal in the output market in the $(I, O$ strategy. This result is attributed to the fact that in the (I, O) strategy, sale are conducted after the output is produced. Hence, using an output pri signal, as opposed to a quantity signal, in the conduct of sales is disadva tageous since it may incur a loss due to unsold inventories whenever output demand falls below the production level.

The only situation where the price signal dominates the quantity signal is in the (O, I) strategy whenever the expected marginal cost is a decreasing function of output (or concave expected cost function). To explain this result, again observe that the concavity of the expected cost function implies, by Jensen's inequality, that expected cost is smaller (and hence the expected profit is greater), ceteris paribus, the greater the uncertainty in quantities of output demand. Thus, because the price, but not the quantity, signal elicits uncertain quantity responses in output demand, it follows that the output price signal is superior.

Thus, except for the singular case of decreasing expected marginal costs in the (O, I) behavioral class, quantity generally dominates price signals under most plausible conditions. This is a striking result which have important implications. In a world without an omniscient auctioneer who freely supplies information on trading prices, price setting by the firm is a very natural way in the conduct of trade. In many cases, its dominance by strategies involving quantity signals is indeed surprising. To rationalize this, observe that a quantity signalling experiment presumes that for any offer of a quantity for sale or purchase, there exists some response price at which no more than whatever quantity is offered by the firm is bought or sold. The natural question, then, is how such a response is being determined. This question is not explicitly considered in our model in which a price response to a quantity signal is treated symmetrically as an exogenous variable just as is a quantity response to a price signal. However, it is clear that the feasibility of a quantity signalling experiment may have to depend on the existence of intermediate trading agents that would respond with a trading price to the firm's quantity offer. Interpreted in this manner, the very domination of a quantity over a price signalling experiment suggests that firms should be willing, in order to operate a quantity signalling experiment, to offer premiums to intermediate trading agents rather than having to use price signals to conduct output sales (retailing) or input purchase where it directly assumes the risk of unsold inventory or under-employment of input, respectively. This risk is effectively shifted to intermediate trading agents in a quantity signalling experiment. The existence of private employment agencies, and retailers are some examples of the intermediate trading agents discussed above. Their existence is otherwise difficult to rationalize in a model without uncertainty since the firm could very well have undertaken all the input purchase and output selling activities by itself.

The *second issue* we addressed was the relative dominance between experiments that belong to different behavioral classes differentiated according to the sequence of experimentation between the input and output markets. On this matter, and abstracting from the issue of price versus quantity decisions, we found that the $(n^d, q^s) \in (I, O)$ strategy to be most preferred if the input market uncertainty is large relative to the output market uncertainty, but the $(q^s, n^d) \in (O, I)$ strategy is most preferred if the converse is true. This fairly obvious result is explained below. In $(n^d, q^s) \in (I, O)$ or "buy-then-sell" strategy, the firm experiments first in the input market and any further uncertainty in this market does not exist after it makes its input purchase contract. However, its output sales must bear the remaining uncertainty in the output market. In contrast, the $(q^s, n^d) \in (O, I)$ or "sell-then-buy" strategy, the initial experiment in the output market resolves any further uncertainty in this market after its output sales contract is made. However, its input purchase must bear the brunt of the input market uncertainty. Hence, everything else constant, a strategy that experiments first in the market with a larger degree of uncertainty is generally preferred to one which experiments in it last because the former (as opposed to the latter) strategy would leave a smaller amount of uncertainty to be borne in the remaining market. It should be emphasized that the above explanation for dominance among the different behavioral classes must be qualified by the ceteris paribus condition that the profit functions do not differ among the behavioral modes considered. While this condition holds true for the group of behavior: (n^d, q^s), $\binom{n^d}{q^s}$, and (q^s, n^d), the same cannot be said for the group: (n^d, p^s), $\binom{n^d}{p^s}$, and (p^s, n^d).

Finally, the generalized versions (to include the input market uncertainty) of the three behavioral modes considered by Baron [2], Leland [4], and Lim [6] are but a small subset of the possible behavioral modes considered in this model. In the framework of this model, these behavioral modes not only belong to different behavioral classes but they differ also in the use of price and quantity signals. The evaluation of their relative dominance essentially generalizes Lim's [6] results.

ACKNOWLEDGMENTS

This paper is a considerably revised version of Chapter 3 of the author's Ph.D. thesis submitted to Queen's University. The author is indebted to his thesis committee, Richard Harris and Frank Flatters, and also to Douglas Purvis, Hajime Oniki, Lloyd Paquin, and Dan Usher for their encouragement during the progress of the thesis. The author is also very grateful to an anonymous referee for providing many valuable comments for improvement on the previous drafts.

REFERENCES

1. K. J. Arrow and R. A. Lind, Uncertainty and the evaluation of public investment decisions, *Amer. Econ. Rev.* **60** (1970), 364–378.
2. D. Baron, Demand uncertainty in imperfect competition, *Internat. Econ. Rev.* **12** (1971), 196–208.

3. K. IWAI, The firm in uncertain markets and its price, wage and employment adjustments, *Rev. Econ. Studies* **41** (1974), 257–276.
4. H. LELAND, Theory of the firm facing uncertain demand, *Amer. Econ. Rev.* **62** (1972), 278–291.
5. C. LIM, "Uncertainty and Implicit Contracts," Ph. D. dissertation, Queen's University, February, 1978.
6. C. LIM, The ranking of behavioral modes of the firm facing uncertain demand, *Amer. Econ. Rev.* **70** (1980), 217–224.
7. J. MARSCHAK AND K. MIYASAWA, Economic comparability of information systems, *Internat. Econ. Rev.* **9** (1968), 137–174.
8. E. S. MILLS, Uncertainty and price theory, *Quart. J. Econ.* **73** (1959), 116–129.
9. E. S. PHELPS, *et al.*, "Microeconomic Foundations of Employment and Inflation," Norton, New York, 1970.
10. H. RAIFFA AND R. SCHLAIFFER, "Applied Statistical Decision Theory," Harvard Univ. Press, Cambridge, 1961.
11. P. A. SAMUELSON, The fundamental approximation theorem of portfolio analysis in terms of means, variance, and higher moments, *Rev. Econ. Studies* **37** (1970), 537–542.
12. G. J. STIGLER, The economics of information, *J. Pol. Econ.* **69** (1961), 213–225.
13. M. L. WEITZMAN, Prices versus quantities, *Rev. Econ. Studies* **41** (1974), 479–491.

JOURNAL OF ECONOMIC THEORY 24, 362–376 (1981)

The Theory of Search for Several Goods

KENNETH BURDETT*

Department of Economics, University of Wisconsin, Madison, Wisconsin 53706

AND

DAVID A. MALUEG

Department of Economics, Northwestern University, Evanston, Illinois 60201

The problem analyzed in this study can be illustrated by the following example. Before starting out to buy groceries for the week, suppose an individual makes a list of the n goods required ($n \geqslant 2$). The list also states the amount required of each good. There are many grocery stores the individual can visit. Each of these stores sells all the goods desired. The individual can visit any store at a cost and observe the n vector of prices offered, one price for each of the n goods. The individual can purchase any number of the goods required from this store and then continue to search for the remaining goods. Given each price vector observed can be envisaged as a random draw from a known nondegenerate distribution of price vectors, the problem is to determine the search strategy that minimizes the individual's expected total cost of purchasing the n goods.

The nature of the above problem is similar to that posed in the literature on consumer search for the lowest price for a single good.[1] In the literature on single good search a consumer is assumed to be shopping for a given amount of a good. On the payment of a fixed cost the consumer may visit a store and observe the price offered, which is assumed to be a random draw from a known nondegenerate distribution of prices.[2] Many studies have assumed that the consumer can return to any store preciously visited without incurring a cost; the other studies have assumed that the consumer cannot

---

* Supported in part by an N.S.F. grant.

[1] Lippman and McCall [6] provide an excellent survey of this literature and the related literature on worker job search. Stigler [11] presents an early example of work in this area. Pratt *et al.* [8] have presented evidence that the distribution of prices for many goods is nondegenerate, at least in Boston.

[2] Rothschild [9] considers a model of consumer search where the consumer does not know the distribution of prices in the market. In this case the well-known reservation price property may not hold.

0022-0531/81/030362-15$02.00/0

[illegible] Press, Inc.

return to any store.[3] Which of these two assumptions is used is of no great importance in consumer search for a single good, however, as it has been shown many times that the search strategy that minimizes the consumer's expected cost of purchasing the good is the same in either case. Specifically, it has been shown that the optimal strategy can be characterized by a reservation price, p^*. The consumer should purchase the first time a price no greater than p^* is observed, where

$$c = \int_0^{p^*} (p^* - s)\, dF(s),$$

c is the cost of a price observation, and $F(\cdot)$ is the distribution function associated with any price observation.[4]

In the present study a model similar in most respects to those presented in the single good search literature will be used. There are, however, two important differences. First, it will be assumed that the consumer wants to purchase n goods, where n is any positive integer. Second, the consumer can visit a store and observe the price quotation for each one of the n goods desired. Unlike the single good consumer search models, the assumption made about the consumer's ability to return to any store is critical in consumer search models for several goods. In Section 1 it will be assumed that the consumer may return to any store without incurring a cost. In Section 2 the consumer is assumed to be unable to return to any store. It will be shown that the strategies that minimize the consumer's expected total cost of purchasing differ significantly in the two situations considered, at least in the case where $n = 2$.

Suppose the consumer can freely return to any store previously visited. It follows immediately that in this case the consumer need not purchase any of the goods until he or she selects to by all the n goods. When the decision to purchase all the goods has been made, the consumer can freely return to the store that offered the lowest price observed for good i and purchase good i from that store, $i = 1, 2, \ldots, n$. Using the results presented in the single good consumer search literature, it appears reasonable to conjecture that the consumer's best strategy can be characterized by what may be loosely termed a reservation frontier. In Section 1 this conjecture is shown to be true. Specifically, it is shown that the cost minimizing search strategy implies the consumer should stop searching and purchase all n goods if and only if the vector of lowest prices observed to date is no greater than some point on the reservation frontier. Most of Section 1 is taken up characterizing the reservation frontier and hence the set of acceptable price vectors. It will be demonstrated that the set of acceptable price vectors is convex. Further, it will be shown that the consumer will never purchase good i at a price higher than the reservation price for good i in a single good search for good i, $i = 1, 2, \ldots, n$.

---

[3] To the authors' knowledge, all but one of the studies presented to date have utilized one of these two restrictions. Landsberger and Peled [5] provide the exception.

[4] See Lippman and McCall [6] for a derivation of this result. Note that p^* is independent of the recall assumption.

In the final part of Section 1 we investigate the consequences of changes in the parameters of the model of consumer search. For example, it is shown that an increase in the cost of search will shift out the relevant reservation frontier and hence enlarge the set of acceptable price vectors. The effect of a change in the distribution of prices for one of the goods is also investigated. First, it is shown that if the distribution of one of the n prices is translated up, the set of acceptable price vectors increases. Second, if the distribution of one of the prices becomes riskier in the sense specified by Rothschild and Stiglitz [10], the set of acceptable price vectors decreases.

In Section 2 the case where the consumer cannot return to any store is considered. In this case if the consumer decides to leave a store and visit another one without purchasing, then the consumer loses the opportunity to purchase from that store at the offered price. Some of the prices offered, however, may be too good to miss. In Section 2 we consider a model of this type where the consumer is searching for only two goods, i.e., $n = 2$. In this case it is shown that there exists a reservation frontier as in the recall case. The nature of this frontier, however, is different. The strategy that minimizes the expected cost of purchasing when the consumer cannot return to stores can be briefly summerized as follows. First, the consumer will not purchase good i from a store that offers a price for good i that is greater than p_i^*, where p_i^* is the reservation price the consumer would use in a single good search for good i, $i = 1, 2$. Second, given good i has not been purchased to date, the consumer will always purchase good i from a store that offers a price for good i no greater than $\hat{p}_i$, $\hat{p}_i < p_i^*, i = 1, 2$. Finally, suppose the consumer has purchased neither of the two goods to date. If the consumer observes a price for good i, say p_i, such that $\hat{p}_i < p_i < p_i^*$, then both of the goods will be purchased if the price of the other good is no greater than some price $h_i(p_i)$, where $h_i(p_i)$ is a decreasing function, $i = 1, 2$.

It should be stressed that in either of the models considered in this study the consumer need not purchase all the goods desired from the same store. Suppose, for the moment, that the consumer must purchase all the n goods from the same store. In this case the consumer search for several goods is formally equivalent to a single good search model. Specifically, the consumer will select a reservation "price," z^*, and then purchase from the first store visited which offers a price vector $(p_1, p_2, \ldots, p_n)$ such that $\sum_{i=1}^{n} p_i \leqslant z^*$.

1. THE MODEL

Suppose a consumer wants to purchase n different goods ($n \geqslant 2$). One unit of each of these n goods is desired. There are many stores that sell all of these goods. By paying c ($c > 0$) the consumer can visit a store and observe the vector of prices offered, one price for each one of the n goods. The stores the consumer can visit, however, do not all offer the same price vector. Specifically, assume there is a known nondegenerate distribution function, G, that describes the vectors offered by these stores. Assume that each price vector observed can be envisaged as a random draw from this distribution. After visiting a store the consumer may purchase some number of the goods desired and/or visit another store at cost c. Although the consumer can purchase only from a store visited, all of the n goods need not be purchased from the same store.

Given the environment specified above, the problem is to determine and characterize the search strategy that minimizes the consumer's expected total cost of purchasing the n goods. Throughout this section it will be assumed that the consumer can return to any store previously visited without incurring a cost. In the next section the consequences of utilizing an alternative to this free recall assumption will be considered.

Let p_i denote the lowest price for good i observed by the consumer to date, $i = 1, 2,..., n$; and let $p = (p_1, p_2,..., p_n)$ be the vector of these prices. If the consumer has observed at least one price vector, there are three options open.

Option 1. Purchase all n goods and thus end the search process.

Option 2. Purchase none of the goods and then observe another price vector.

Option 3. Purchase fewer than n goods and observe another price vector.

If the consumer is starting the search process and has not observed a price vector, then Option 2, must be selected. The free recall assumption guarantees that the expected cost of purchasing when Option 3 is selected is at least as great as the expected cost if Option 2 were selected instead. Hence, without loss of generality, assume that Option 1 or 2 only will be chosen. Specifically, Option 2 will be chosen at any time until Option 1 is chosen. The search process then ends.

The free recall assumption implies that $p \cdot 1 = \sum_{i=1}^{n} p_i$ denotes the minimum cost of purchasing when Option 1 is chosen and p is the vector of lowest prices observed to date. The consumer will choose Option 1 at any time if and only if $p \cdot 1$ is no greater than the expected cost of purchasing if Option 2 were selected instead. When the consumer selects Option 2, however, the expected cost of purchasing depends on the future search strategy to be used.

At any time the consumer is assumed to know the vector of lowest prices observed to date and the probability laws generating the price vectors that may be observed in the future. In what follows it will be shown that any search strategy can be characterized by a stopping rule. To construct any such rule some notation needs to be developed. Let $\Omega = \times_{i=1}^{\infty} R_{+}^{n}$ denote the set of infinite sequences of n vectors consisting of positive reals. Since the distribution function G is assumed to be known, the probability the consumer would observe any measurable set in Ω if he or she were to search indefinitely is well defined. Let (Ω, β, μ) denote the probability measure space, where μ is the probability measure induced by G, and β is the σ-field. The σ-field generated by the first k searches is denoted by β_k, $k = 1, 2, 3, \ldots$. Using the above notation, a stopping rule is an integer valued function $s\colon \Omega \to I$ such that $s^k \in \beta_k$ for all $k = 1, 2, 3, \ldots$, where $s^k = \{\omega \in \Omega \mid s(\omega) = k\}$. It is straightforward to check that any search strategy based on the information known to the consumer can be described by a stopping rule. For example, suppose the consumer utilizes the following simple search strategy: observe four price vectors and then purchase. The stopping rule $\bar{s}$ that describes this strategy is such that $\bar{s}^4 = \Omega$, and $\bar{s}^k = \varnothing$ for all $k \neq 4$. Let $\mathcal{S}$ denote the set of all possible stopping rules.

Suppose the consumer has decided to observe at least one more price vector when p denotes the vector of lowest prices observed to date. If the consumer utilizes the search strategy that can be described by stopping rule s, the expected total cost by purchasing, $\psi(p, s)$, can be written as

$$\psi(p, s) = \sum_{k=1}^{\infty} \int_{s^k} \left\{ \sum_{i=1}^{n} \min(p_i, W_{ik}) + ck \right\} d\mu(\omega), \tag{1}$$

where W_{ik} is the random variable describing the lowest price observed for good i in k observations.[5] In (1) W_{ik} is conditioned by $\omega \in s^k$. The expected cost of purchasing when any search strategy is used depends on the probability measure μ which is induced by G. Throughout the study it will be assumed that G has a compact support. Further, let $V = \{p \in R_{+}^{n} \mid \underline{p} \leqslant p \leqslant \bar{p}\}$ denote the smallest closed n rectangle enclosing the support, i.e.,

$$\begin{aligned} G(p) &= 0 && \text{for all } p < \underline{p}, \\ 0 < G(p) &< 1 && \text{for all } \underline{p} \leqslant p < \bar{p}, \\ G(p) &= 1 && \text{for all } p \geqslant \bar{p}.^{6} \end{aligned}$$

---

[5] Specifically, $W_{ik} = \min(P_{i1}, P_{i2}, \ldots, P_{ik})$, where P_{ij} denotes the random variable describing the price the consumer may observe for good i in the jth search, $i = 1, 2, \ldots, n$.

[6] Throughout this study it will be assumed that $(p_1, p_2, \ldots, p_n) \geqslant (p'_1, p'_2, \ldots, p'_n)$ if and only if $p_i \geqslant p'_i$, $i = 1, 2, \ldots, n$. Further, the inequality is strict of $p_i \geqslant p'_i$ for $i = 1, 2, \ldots, n$ and $p_j > p'_j$ for some $j = 1, 2, \ldots, n$.

Terming $\underline{p}_i$ ($\bar{p}_i$) the lowest possible (highest possible) price for good i, although formally incorrect, should lead to no confusion. It is possible to weaken this restriction on G without disturbing the basic results to be presented. The restriction used, however, appears natural in the situation envisaged. Below an important result is stated. As the proof of this claim has been presented many times before (see, for example, DeGroot [2, Chap. 13, Sec. 8]), none will be presented.

Claim 1. Given G has a compact support, for any fixed $p \in R^n_+$ there exists an s^*_p such that $\psi(p, s^*_p) \leqslant \psi(p, s)$ for any $s \in \mathscr{S}$.

The stopping rule s^*_p is termed the optimal stopping rule when p is the vector of lowest prices observed to date. Utilizing Claim 1 we may define

$$\phi(p) = \min_{s \in \mathscr{S}} \psi(p, s) \tag{2}$$

to be the expected cost of purchasing when Option 2 is chosen.

Above it has been shown that the expected cost of purchasing is well defined when either Option 1 or Option 2 is chosen. In either case the expected cost of purchasing depends on the vector of lowest prices observed to date. When will Option 1 be chosen, i.e., when will the consumer decide to purchase the n goods? In what follows the answer to this question will be presented. Specifically, the set of acceptable prices $A = \{p \in R^n_+ \mid p \cdot 1 \leqslant \phi(p)\}$ will be characterized. To achieve such a goal three lemmas are first stated and proved.

LEMMA 1. *$\phi(\cdot)$ is a concave nondecreasing function.*

Proof. To establish concavity it is shown that

$$\phi(\alpha p + (1-\alpha)\tilde{p}) - \alpha\phi(p) - (1-\alpha)\,\phi(\tilde{p}) \geqslant 0,$$

for any p, $\tilde{p}$, and α, where $0 \leqslant \alpha \leqslant 1$. Let $\hat{s}$ be the optimal stopping rule when $(\alpha p + (1-\alpha)\tilde{p})$ denotes the vector of lowest prices observed to date, i.e., $\phi(\alpha p + (1-\alpha)\tilde{p}) = \psi(\alpha p + (1-\alpha)\tilde{p}, \hat{s})$. Since (2) guarantees that $\psi(p, \hat{s}) \geqslant \phi(p)$ and $\psi(\tilde{p}, \hat{s}) \geqslant \phi(\tilde{p})$, it follows that

$$\begin{aligned}
&\phi(\alpha p + (1-\alpha)\tilde{p}) - \alpha\phi(p) - (1-\alpha)\,\phi(\tilde{p}) \\
&\quad \geqslant \psi(\alpha p + (1-\alpha)\tilde{p}, \hat{s}) - \alpha\psi(p, \hat{s}) - (1-\alpha)\,\psi(\tilde{p}, \hat{s}) \\
&\quad = \sum_{k=1}^{\infty} \int_{\hat{S}^k} \sum_{i=1}^{n} \{\min(\alpha p_i + (1-\alpha)\tilde{p}_i, W_{ik}) \\
&\qquad - \alpha \min(p_i, W_{ik}) - (1-\alpha)\min(\tilde{p}_i, W_{ik})\}\, d\mu(\omega).
\end{aligned} \tag{3}$$

Clearly,

$$\min(ap_i + (1-a)\bar{p}_i, W_{ik}) - a\min(p_i, W_{ik}) - (1-a)\min(\bar{p}_i, W_{ik}) \geqslant 0 \quad (4)$$

for any i and k. Inspection of (3) and (4) establishes the concavity of $\phi(\cdot)$. To demonstrate that $\phi(\cdot)$ is nondecreasing note that for any $s \in \mathscr{S}$ $\psi(p, s) \geqslant \psi(p', s)$, if $p > p'$. This fact and (2) imply $\phi(p) \geqslant \phi(p')$, if $p > p'$. This completes the proof of the lemma.

Using Lemma 1 it follows immediately that $(\phi(p) - p \cdot 1)$ is a concave function of p and thus the set of acceptable prices A is convex. To facilitate the exposition let G_k denote the marginal distribution for the price of good k, $k = 1, 2, 3,\ldots, n$, i.e.,

$$G_k(p_k) = \int_{(\underline{p}_1,\ldots,\underline{p}_{k-1},\underline{p}_k,\underline{p}_{k+1},\ldots,\underline{p}_n)}^{(\bar{p}_1,\ldots,\bar{p}_{k-1},p_k,\bar{p}_{k+1},\ldots,\bar{p}_n)} dG(p)$$

for any p_k.

Lemma 2. *Suppose $p \in R^n_+$ is such that $p_i \leqslant \underline{p}_i$ for $i = 1, 2,\ldots, n$, $i \neq k$, where k is an integer such that $1 \leqslant k \leqslant n$. In this case $\phi(p) \gtrless p \cdot 1$ as $p_k \lesseqgtr p_k^*$, where*

$$c = \int_{\underline{p}_k}^{p_k^*} (p_k^* - s)\, dG_k(s). \quad (5)$$

Proof. If p satisfies the hypothesis of the lemma, then with probability one no lower price will be observed for good i, $i \neq k$. Hence, letting $\Pi(p) = \min(p \cdot 1, \phi(p))$, we have

$$\phi(p) = c + \int_{\underline{p}_k}^{p_k} \Pi(p_1,\ldots, p_{k-1}, s, p_{k+1},\ldots, p_n)\, dG_k(s) + [1 - G_k(p_k)]\, \Pi(p),$$

if the kth element of p is p_k and $p_i \leqslant \underline{p}_i$, $i \neq k$. Rearranging terms yields

$$\phi(p) - \Pi(p) = c - \int_{\underline{p}_k}^{p_k} [\Pi(p) - \Pi(p_1,\ldots, p_{k-1}, s, p_{k+1},\ldots, p_n)]\, dG_k(s).$$

Since $\phi(p) - \Pi(p) = c > 0$ when $p_i \leqslant \underline{p}_i$, $i = 1, 2,\ldots, n$, $\Pi(p) = p \cdot 1$ in some neighborhood of $\underline{p} = (\underline{p}_1, \underline{p}_2,\ldots, \underline{p}_n)$. Consequently, for any p satisfying the hypothesis of the lemma,

$$\begin{aligned} \phi(p) - \Pi(p) &= c - \int_{\underline{p}_k}^{p_k} (p_k - s)\, dG_k(s) > 0, && \text{if } p_k < p_k^*, \\ &= 0, && \text{if } p_k \geqslant p_k^*. \end{aligned}$$

This completes the proof of the lemma.

Lemma 2 links the search for several goods with the theory of search for one good. Specifically, if the consumer has observed the lowest possible price for $n-1$ goods, then the reservation price used for the other good is precisely that which would be used in a single good search for that good; i.e., p_k^* is the reservation price the consumer would use if searching for a low price for good k only, $k = 1, 2,\ldots, n$. The next lemma demonstrates that if the consumer prefers to purchase all n goods when p is the vector of lowest prices observed to date, then the consumer would also prefer to purchase if $p' < p$ is the vector of lowest prices observed to date.

LEMMA 3. *If* $\phi(p) = p \cdot 1$, *then*

(i) $p^1 < p$ *implies* $\phi(p^1) \geqslant p^1 \cdot 1$, *with inequality strict if* $p_i > p_i^1$ *and* $G_i(p_i) > 0$ *for some* i, $i = 1, 2,\ldots, n$, *and*

(ii) $p^2 > p$ *implies* $\phi(p^2) \leqslant p^2 \cdot 1$, *with the inequality strict if* $p_i^2 \neq p_i$ *and* $G_i(p_i^2) > 0$ *for some* i, $i = 1, 2,\ldots, n$.

Proof. To establish the claims of the lemma it is shown that for any two vectors p^1 and p^2 such that $p^2 > p^1$, $\phi(p^2) - \phi(p^1) \leqslant p^2 \cdot 1 - p^1 \cdot 1$, with the inequality strict if $p_i^1 \neq p_i^2$ and $G_i(p_i^2) > 0$ for some $i = 1, 2,\ldots, n$. Let s_1 denote the optimal stopping rule when p^1 is the vector of lowest prices observed to date, i.e., $\phi(p^1) = \psi(p^1, s_1)$. From (2) it follows that $\phi(p^2) \leqslant \psi(p^2, s_1)$. Hence,

$$\begin{aligned}
\phi(p^2) - \phi(p^1) &\leqslant \psi(p^2, s_1) - \psi(p^1, s_1) \\
&= \sum_{k=1}^{\infty} \int_{s_1^k} \sum_{i=1}^{n} \{\min(p_i^2, W_{ik}) - \min(p_i^1, W_{ik})\}\, d\mu(\omega) \\
&\leqslant \sum_{i=1}^{n} (p_i^2 - p_i^1).
\end{aligned}$$

Note that the last inequality is strict if $p_i^2 > p_i^1$ and $G_i(p_i^2) > 0$ for some i, $i = 1, 2,\ldots, n$. The two claims made in the lemma now follow directly from the result established above.

The lemmas established above are now used to prove the main result of this section. The theorem provides a complete characterization of the optimal strategy of a consumer searching for n goods with free recall.

THEOREM 1. (i) *The set of acceptable prices* $A = \{p \in R_+^n \mid \phi(p) \geqslant p \cdot 1\}$ *is convex.*

(ii) *The set* A *is a strict subset of the set* $T = \{p \in R_+^n \mid p_i \leqslant p_i^*, i = 1, 2,\ldots, n\}$.

(iii) *If* $\phi(p) = p \cdot 1$, *then*

$$c = \sum_{i=1}^{n} \int_{\underline{p}_i}^{p_i} (p_i - s)\, dG_i(s). \tag{6}$$

Further,

$$\left.\frac{dp_j}{dp_i}\right|_{\phi(p) = p \cdot 1} = -\frac{G_i(p_i)}{G_j(p_j)}, \qquad \textit{for } i \neq j,\ i = 1, 2, \ldots, n \textit{ and } j = 1, 2, \ldots, n.$$

Proof. (i) From Lemma 1 it follows that $(\phi(p) - p \cdot 1)$ is a concave function of p. Hence, A is a convex set.

(ii) Lemma 2 and Lemma 3 establish that if $p \in A$, then $p_i \leqslant p_i^*$, $i = 1, 2, \ldots, n$. This implies $A \subset T$. Lemmas 2 and 3 also imply that $(p_1^*, p_2^*, \ldots, p_n^*) \notin A$. Thus $T \not\subset A$.

(iii) Assume that if p denotes the vector of lowest prices observed to date, then $\phi(p) = p \cdot 1$. Using Lemma 3 it follows that

$$\begin{aligned}
\phi(p) &= c + E\left\{\sum_{i=1}^{n} \min(p_i, W_{i1})\right\} \\
&= c + \Pr(W_{n1} \geqslant p_n)\left[p_n + E\left\{\sum_{i=1}^{n-1} \min(p_i, W_{i1}) \mid W_{n1} \geqslant p_n\right\}\right] \\
&\quad + \Pr(W_{n1} < p_n)\left[E(W_{n1} \mid W_{n1} < p_n)\right. \\
&\quad \left. + E\left\{\sum_{i=1}^{n-1} \min(p_i, W_{n1}) \mid W_{n1} < p_n\right\}\right] \\
&= c + E\left\{\sum_{i=1}^{n-1} \min(p_i, W_{i1})\right\} - \int_{\underline{p}_n}^{p_n} (p_n - s)\, dG_n(s) + p_n.
\end{aligned}$$

Continuing in the same fashion we obtain the following relationship for any k, $k = 1, 2, \ldots, n$:

$$\phi(p) = c + E\left\{\sum_{i=1}^{k-1} \min(p_i, W_{i1})\right\} - \sum_{i=k}^{n} \int_{\underline{p}_i}^{p_i} (p_i - s)\, dG_i(s) + \sum_{i=k}^{n} p_i.$$

Since $\phi(p) = p \cdot 1$ by assumption, $\phi(p) = \sum_{i=1}^{n} p_i$. This establishes the first part of claim (iii). To prove the second part of claim (iii) note that integrating (6) by parts yields

$$\sum_{i=1}^{n} \int_{\underline{p}_i}^{p_i} (p_i - s)\, dG_i(s) = \sum_{i=1}^{n} \int_{\underline{p}_i}^{p_i} G_i(s)\, ds. \tag{7}$$

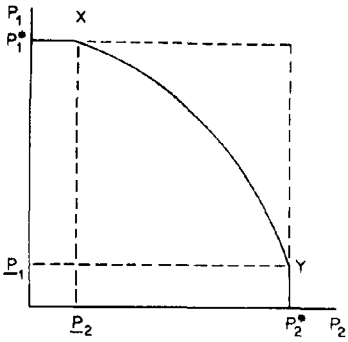


FIGURE 1

To illustrate the results presented above we consider the case where the consumer is searching for two goods, i.e., $n = 2$. This case is illustrated in Fig. 1. The boundary of the acceptance set where $\phi(p_1, p_2) = p_1 + p_2$ is given by the curve $p_1^*XYp_2^*$. Note that this curve in Fig. 1 is horizontal in the section p_1^*X and vertical in the section Yp_2^*. The curve is drawn this way to represent the fact $\phi(p_1^*, p_2) = p_1^* + p_2$ if $p_2 \leqslant \underline{p}_2$, and $\phi(p_1, p_2^*) = p_1 + p_2^*$ if $p_1 \leqslant \underline{p}_1$.[7]

Suppose there is an increase in the cost of search, c. From (7) it follows that this increase will lead to an increase in the set of acceptable price vectors. Specifically, if the price vector p is such that $\phi(p) = p \cdot 1$ when p is the vector of lowest prices observed to date and the cost of search is c, then $\phi(p) > p \cdot 1$ when the cost of search is $c' > c$. The result is illustrated in Fig. 2a for the case where $n = 2$. In Fig. 2a an increase in the cost of search shifts out the reservation frontier from the curve $p_1^*p_2^*$ to the curve $\bar{p}_1^*\bar{p}_2^*$.

In the final part of this section the consequences of a change in the distribution of one of the prices is discussed. There are many ways this distribution can change. Two special cases are considered. First, assume there is an upward translation in the mean of the distribution of prices of good k, $1 \leqslant k \leqslant n$; i.e., the new distribution of prices for good k, $G_k'(\cdot)$, is such that for some $\varepsilon > 0$, $G_k(p_k) = G_k'(p_k + \varepsilon)$ for any p_k. Second, we consider the case where the new distribution of prices for good k is less risky than the old one. Following Rothschild and Stiglitz [10], we define the distribution function $G_k''(\cdot)$ to be less risky than $G_k(\cdot)$, which has the same mean, if

$$\int_{\underline{p}_k}^{z} (G_k(s) - G_k''(s))\,ds \geqslant 0$$

---

[7] The frontier of the acceptance set in Figure 1 has been drawn as if the function $G(\cdot)$ is strictly increasing on its support. If $G(\cdot)$ is not strictly increasing on its support, the frontier will have "flat" sections even if $p_i > \underline{p}_i$, $i = 1, 2$.

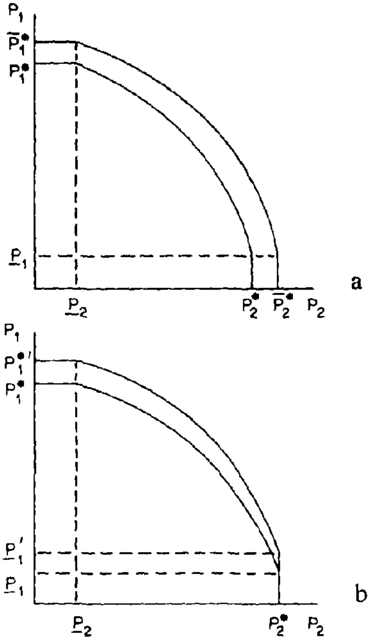


FIGURE 2

for any z, and as $z \to \infty$ the above integral tends to zero. Using (7) it follows immediately that either of the two changes in the distribution function, $G_k(\cdot)$, will increase the set of acceptable price vectors. Specifically, given either of the two changes, if $\phi(p) = p \cdot 1$ when p is the vector of lowest prices observed to date, then an increase in the mean or a reduction in the riskiness of the distribution function $G_k(\cdot)$ will imply $\phi(p) \geqslant p \cdot 1$, with the inequality strict if and only if $p_k > \underline{p}_k$. An increase in the mean of the distribution of prices of good 1 is illustrated in Fig. 2b for the case where $n = 2$. In Fig. 2b the upward translation in the mean has increased the lowest possible price from $\mathbf{p}_1$ to $\mathbf{p}_1'$. The old reservation frontier is illustrated by the curve $p_1^* p_2^*$, the new one by the curve $p_1^{*\prime} p_2^*$.

2. THE NO RECALL MODEL

In the previous section it was assumed that the consumer could return to any store previously visited without incurring a cost. In this section this restriction is dropped. Instead, it will be assumed that the consumer cannot return to any store previously visited. It will be shown that this new assumption leads to a drastic change in the set of acceptable price vectors at least in the case where $n = 2$. It should be noted that the "free recall" assumption used in the previous section and the "no recall" assumption used in this section do not exhaust all possibilities. For example, it could be assumed that the consumer can return freely to the last four stores visited only. Nevertheless, the vast majority of contributions to the search literature have used one of the two restrictions used in the present study.

Suppose the consumer visits a store and observes the vector of prices offered. Since this consumer cannot return to this store in the future, some of the prices offered may be too good to miss and therefore the consumer may purchase some subset of the goods and then continue to search for the other goods. As there are 2^n subsets of a set of cardinality n, there are 2^n options open to the consumer when visiting a store. Computing the expected total cost of purchasing the n goods when each of these options is selected is tedious when n is large, and leads to no new insights. Hence, to illustrate the essentials, the special case where $n = 2$ will be considered in this section.

Suppose the consumer visits a store and observes price p_1 for good 1 and price p_2 for good 2. There are four options open to the consumer.

Option A. Buy both goods and thus end the search process.

Option B. Buy good 2 and then continue to search for a low price for good 1.

Option C. Buy good 1 and then continue to search for a low price for good 2.

Option D. Buy neither of the goods and thus continue to search for both goods.

If Option A is chosen, then the cost of purchasing is $p_1 + p_2$. If Option B is chosen, then $p_1 + V_2$ is the expected cost of purchasing both goods, where V_2 is the minimum expected cost of purchasing good 2 in a single good search. Similarly, $p_2 + V_1$ is the expected cost of purchasing both goods when Option C is chosen. Let W denote the expected cost of purchasing when Option D is selected. Note that the no recall assumption implies W, V_1, and V_2 are independent of the prices observed to date. The consumer will choose the option that minimizes the expected cost of purchasing. Hence, the minimum expected cost of purchasing when prices p_1 and p_2 have just been observed can be written as

$$\min(p_1 + p_2, p_1 + V_2, p_2 + V_1, W). \tag{8}$$

It is well known that the expected cost of purchasing good i in a single search when at least one more price observation is to be taken, V_i, can be expressed as

$$V_i = c + \int_{\underline{p}_i}^{p_i^*} s\, dG_i(s) + (1 - G_i(p_i^*))\, V_i, \tag{9}$$

where $p_i^* = V_i$, $i = 1, 2$ (Lippman and McCall [6] provide a derivation of this result). Note that the reservation price, p_i^*, in the no recall case is the same as the reservation price for good i in the free recall case, other things being equal.

Assume the consumer selects Option D. The expected cost of purchasing will depend on the future search strategy to be used. Suppose the consumer uses the following strategy: visit another store and purchase good i if and only if the price observed is no greater than p_i^*, $i = 1, 2$. If the consumer has not purchased good i after this observation, then he or she will search in an optimal manner for good i in a single good search. When this strategy is used the expected cost of purchasing, $\bar{W}$, can be written as

$$\begin{aligned}
\bar{W} = c &+ \Pr(p_1 \leqslant p_1^* \text{ and } p_2 \leqslant p_2^*)\, E\{p_1 + p_2 \mid p_1 \leqslant p_1^* \text{ and } p_2 \leqslant p_2^*\} \\
&+ \Pr(p_1 \leqslant p_1^* \text{ and } p_2 > p_2^*)[E\{p_1 \mid p_1 \leqslant p_1^* \text{ and } p_2 > p_2^*\} + V_2] \\
&+ \Pr(p_1 > p_1^* \text{ and } p_2 \leqslant p_2^*)[E\{p_2 \mid p_1 > p_1^* \text{ and } p_2 \leqslant p_2^*\} + V_1] \\
&+ \Pr(p_1 > p_1^* \text{ and } p_2 > p_2^*)(V_1 + V_2) \\
= c &+ \int_{\underline{p}_1}^{p_1^*} s\, dG_1(s) + \int_{\underline{p}_2}^{p_2^*} s\, dG_2(s) + (1 - G_1(p_1^*))\, V_1 + (1 - G_2(p_2^*))\, V_2.
\end{aligned}$$

Using (9) it follows immediately that $\bar{W} < V_1 + V_2$. Hence, as $\bar{W} \geqslant W$, we have $W < V_1 + V_2$. This implies the expected total cost of purchasing both goods is lower in a search for both the goods at once than single good search for these goods. It is straightforward to demonstrate that the consumer will always choose to purchase good i when a price no greater than $\underline{p}_i$ is observed, $i = 1, 2$. Using this result it follows that $\underline{p}_1 + V_2 \leqslant W$ and $\underline{p}_2 + V_1 \leqslant W$, and therefore $V_i \leqslant W$, $i = 1, 2$, with the inequality strict if $\underline{p}_i > 0$. From the above arguments we may state the following result:

$$V_1 + V_2 > W \geqslant V_i, \qquad i = 1, 2. \tag{10}$$

This result is important in characterizing the set of acceptable prices in the no recall case.

Given the consumer has visited a store and observed the prices p_1 and p_2, when will the consumer purchase good 1? By virtue of (8), the consumer will purchase good 1 from the store visited if and only if

$$\min(p_1 + p_2,\, p_1 + V_2) \leqslant \min(p_2 + V_1,\, W). \tag{11}$$

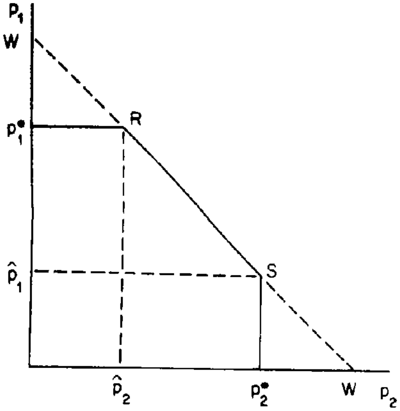


FIGURE 3

Using (10) and (11) it follows that the consumer will purchase good 1, independent of the price offered for good 2, if $p_1 \leqslant \hat{p}_1$, where $\hat{p}_1 + p_2^* = W$. Further, inspection of (10) and (11) establishes that the consumer will never purchase good 1 from a store that offers a price $p_1 > p_1^*$. If the two prices observed are such that $p_i < p_i \leqslant p_i^*$, $i = 1, 2$, the consumer will purcase both goods if and only if $p_1 + p_2 \leqslant W$, otherwise the consumer will continue to search for both goods. This result leads to a complete characterization of the set of acceptable prices in the no recall case when $n = 2$. The conclusions are summarized in Theorem 2 and illustrated in Fig. 3.

THEOREM 2. *Given the consumer has purchased neither of the two goods and has just observed prices p_1 and p_2, the following strategy minimizes the consumer's expected cost of purchasing both goods.*

(a) *Select Option* A *if and only if $p_1 + p_2 \leqslant W$ and $p_i \leqslant p_i^*$, $i = 1, 2$.*

(b) *Select Option* B *if and only if $p_1 \leqslant \hat{p}_1$ and $p_2 > p_2^*$.*

(c) *Select Option* C *if and only if $p_2 \leqslant \hat{p}_2$ and $p_1 > p_1^*$.*

(d) *Select Option* D *if and only of none of the above conditions are satisfied.*

If only one of the goods has been purchased, the consumer should continue to search for the other good as in an optimal single good search.

Let A^* denote the set of acceptable price vectors in the no recall case. A^* is illustrated by the area $0p_1^*RSp_2^*$ in Fig. 3. Using Theorems 1 and 2 it follows that, ceteris paribus, the set of acceptable price vectors in the free recall case, A, is a strict subset of A^*. Thus, there are advantages to the consumer having the right of free recall. This is not true with consumer search for a single good.

ACKNOWLEDGMENTS

We would like to thank C. Wilson, M. Rothschild, D. Mortensen, and the anonymous referee for helpful comments. They are, of course, not responsible for any remaining errors.

REFERENCES

1. *L. Brieman, Stopping rule problems, in "Applied Combinatorial Mathematics" (E. F. Beckenbach, Ed.), Univ. of California Press, Berkeley 1964.*
2. *M. H. DeGroot, "Optimal Statistical Decisions," McGraw-Hill, New York, 1970.*
3. *J. L. Gastwirth, On probabilistic models of consumer search for information, Quart. J Econ.* **90** *(1) (February 1976), 38–50.*
4. M. G. Kohn and S. Shavell, The theory of search, *J. Econ. Theory* 9 (2) (October 1974), 93–123.
5. M. Landsberger and D. Peled, Duration of offers and price realization, *J. Econ. Theory* **16** (January 1977), 17–37.
6. S. A. Lippman and J. J. McCall, The economics of job search: A survey, *Econ. Inquiry* **14** (June 1976), 155–189, and (Sept. 1976), 347–368.
7. J. J. McCall, The economics of information and optimal stopping rules, *J. Business* **38** (July 1965), 300–317.
8. J. Pratt, D. Wise, and R. Zeckhauser, "Price Variations in Almost Competitive Markets," Discussion Paper 37D, Kennedy School of Government, Harvard U., December 1975.
9. M. Rothschild, Searching for the lowest price when the distribution of prices is unknown, *J. Pol. Econ.* (July/August 1974), 689–711.
10. M. Rothschild and J. Stiglitz, Increasing risk I: A definition, *J. Econ. Theory* **2** (1970), 225–243.
11. G. J. Stigler, The economics of information, *J. Pol. Econ.* (June 1961), 213–225.

The Steady States of a Stochastic Decentralized Growth Model

JOHN LAITNER

Department of Economics, University of Michigan,
Ann Arbor, Michigan 48109

Received April 20, 1979; March 24, 1980

The purpose of this paper is to prove the existence of a steady state for a simple decentralized growth model which is subject to random shocks. The model has the following elements. The population is composed of an endless sequence of overlapping cohorts of two-period households; each cohort is the same size; and, there is an aggregate production function and one output good. Because the labor supply is fixed, the production function relates output in each period t to the total saving of all young families in period $t-1$ and to the current realization z_t of a random variable $\tilde{z}$. The random variable has an exogenous, unchanging, and known distribution. Futures markets are imperfect: young families can save at time t only by investing in physical capital, which will pay a return at time $t+1$. The return depends on the size of the time-$(t+1)$ capital stock, k_t, and on z_{t+1}. Let $y_t = y(k_{t-1}, z_t)$ be the vector of incomes of young families at time t. Then such families use a function $g(y_t) = k_t$ to predict k_t (note that g is nonstochastic). The families adjust their saving to maximize their expected utility, where the latter is computed using g and the distribution function for $\tilde{z}$. We call resulting aggregate savings by young families $\gamma(y_t, g)$.

We use a fixed-point argument to prove the existence of sets K and Y and an expectation function g such that if all realizations of $\tilde{z}$ fall within Z, (i) $y(K \times Z) \subset Y$, (ii) $g(Y) \subset K$, and (iii) $g(y) = \gamma(y, g)$ all $y \in Y$. The resulting g determines (in a very simple context) what Grandmont [7] would call "'rational' expectations" and what Radner [17, 18] would call an "equilibrium of plans and expectations." We use this g to generate a transition function v for k_t. We then show v has a finite invariant set K^* (which includes all feasible states except one trivial one); there are a finite number of ergodic sets $E_1, \ldots, E_N \subset K^*$ for v; and there is a unique stationary distribution function D_l associated with each E_l. Each pair (D_l, E_l) constitutes a "steady state" (corresponding to Green and Majumdar's [9] concept of an "equilibrium distribution" and to Grandmont's [7] "long-run equilibrium"): Once the evolution of the economy carries it into E_l some

0022-0531/81/030377-16$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

$l \in \{1, \ldots, N\}$, it will remain in E_l in each subsequent period with probability 1. If the economy has entered E_l, $\int_A D_l(dx)$ gives the long-term frequency of appearance for states in any Borel subset $A \subset E_l$. States in $K^* - (\bigcup_{l=1}^{N} E_l)$ are transitory, leading to entrance into some E_l after a finite number of periods with probability 1.

The novelty of our analysis lies in our ability to characterize v solely on the basis of restrictions on family preference orderings and the aggregate production function.[1] Our work can be viewed as a complement to Mirman's [14, 15] competitive analysis and to Brock and Mirman's [4] optimal growth model. Mirman's papers postulate the existence of an aggregative saving function with various properties; Brock and Mirman derive such a function from a model with a single, centrally controlled maximization process. In this paper, utility maximization on the part of individual families leads to an overall saving function.[2]

The organization of this paper is as follows. Section I presents the assumptions of our model. Section II proves the existence of an equilibrium transition function v (in other words, the existence of a transition function consistent with rational expectations on the part of all families and zero excess demand in each period). Our proof is somewhat involved because we need v to have strong enough properties to establish the existence of a steady state in Section III. A mathematical appendix at the end of the paper supplies proofs for all lemmas and propositions.

I. The Model

We first present our basic assumptions about preferences and production. Then we list several special restrictions needed for particular steps in this paper. At the end of this section we prove two lemmas which we will use to construct an invariant set $[k_L, k_U] \subset (0, \infty)$ for values of the aggregate capital stock.

We assume that at each time t the population of the economy consists of an equal number of young and old families. The number remains fixed over time. Although the population is finite, we assume that it is so large that with no discernible influence on its own well-being each individual family can behave as though its saving decision will not affect the next-period wage rate and rate of return on capital. The families in each cohort are indexed with numbers $j \in \{1, \ldots, J\}$. Each family j has a utility function $u(c_1, c_2, j)$ with c_1

---

[1] See, in particular, Grandmont's comments in Sections 5.1–5.3 of [7] and Radner's comments in Section 4.1 of [18].

[2] Inaba [11] uses a model somewhat in the same vein. However, he does not allow production and he does place direct restrictions on aggregative demand functions.

consumption in the first period of life and c_2 consumption in the second. The family's endowment consists of a supply of "effective" labor units in each period of life, $(l_1(j), l_2(j))$. We assume that the number of families of each type $(u(\cdot); l_1(\cdot), l_2(\cdot))$ remains the same generation after generation. We also assume $l_1(j) > 0$ and $l_2(j) \geqslant 0$ each j.

We want $u\colon R^2 \times \{1,\ldots,J\} \to R$ to have the following properties in every case:

(A1) *For each* $(c_1, c_2) \in (0,\infty) \times (0,\infty)$, u *is twice continuously differentiable in* c_1 *and* c_2, $\lim_{c \to 0} u_1(c, c_2, j) = \lim_{c \to 0} u_2(c_1, c, j) = \infty$, $u_1(c_1, c_2, j) > 0$, $u_2(c_1, c_2, j) > 0$, $u_{12}(c_1, c_2, j) \geqslant 0$, *and* u *is concave in* c_1 *and* c_2.

Any utility function $u\colon R^2 \to R$ of the Bergson class, for example, would satisfy (A1).

Let k_t be the total saving of all young families at time t. Since the total supply of "effective" labor units is unvarying, production next period, q_{t+1}, can be written

$$q_{t+1} = f(k_t, z_{t+1}), \tag{1}$$

where z_{t+1} is a random realization of $\tilde{z}$. This formulation is consistent with the spirit of Mirman's [14, 15] and Brock and Mirman's [4] familiar analyses. We assume that z_t all t represent random samplings from $\tilde{z}$, that $z_t \in R$ all t, and that

(A2) *There exists an interval* $Z = [z_L, z_U] \subset (-\infty, \infty)$ *with* $\int_Z P(dz) = 1$, *where* P *is the distribution function of* $\tilde{z}$.

Since there is only one good available in each period, we can think of its current price as always being 1. Then assuming constant returns to scale and competitive conditions, if w_t is the wage rate per "effective" unit of labor at time t and r_t is one plus the interest rate (net of depreciation),

$$w_{t+1} = w(k_t, z_{t+1}) = f(k_t, z_{t+1}) - f_1(k_t, z_{t+1}) \cdot k_t, \tag{2}$$

$$r_{t+1} = r(k_t, z_{t+1}) = f_1(k_t, z_{t+1}). \tag{3}$$

We also assume

(A3) f *is twice continuously differentiable. For each* $k > 0$ *and* $z \in Z$ *we have* $f(k,z) > 0$, $r(k,z) > 0$, $w(k,z) > 0$, $w_1(k,z) > 0$, *and* $r_1(k,z) < 0$. *We also assume* $\lim_{k \to \infty} r(k,z) < 1$ *and* $\lim_{k \to 0} r(k,z) = \infty$ *any* $z \in Z$.

We need three specialized assumptions. The first will enable us to prove the existence of an ergodic set for k other than $\{0\}$.[3] *If l is the total supply of "effective" labor units per period, let*

$$\sigma(k, z) = r(k, z) \cdot k/[w(k, z) \cdot l].$$

Let "lim" stand for "$\lim_{n \to \infty}$." Then

(A4) For at least one j, given any sequence $\{(a(n), a^*(n))\}$ with $\lim(a(n), a^*(n)) = (0, 0)$, if $\lim u_1(a(n), a^*(n), j)/u_2(a(n), a^*(n), j) = \infty$, we must have $\lim a^*(n)/a(n) = \infty$. Also, $\lim_{k \downarrow 0} \sigma(k, z) \in [0, \infty)$ any $z \in Z$, and $\lim_{k \downarrow 0} w(k, z)/w(k, z') \in (0, \infty)$ any $z, z' \in Z$.

The first part of this assumtion requires that if we graph the isoquant map for u, moving through the positive orthant toward the origin, the marginal rate of substitution must not diverge to ∞ unless the path we follow converges to the $c_2(j)$ axis. Homothetic utility functions, for instance, satisfy this restriction (given (A1)). CES production functions with elasticity of substitution $\geqslant 1$ will satisfy the second part of the assumption.

In order to prove the existence of an equilibrium transition function with several essential properties in Section II, we need

(A5) $\qquad d[r(k, z) \cdot k]/dk > 0 \qquad$ all $k \in K$, $z \in Z$,

where K is a bounded interval to be defined shortly. This assumption ensures that an additional dollar's worth of saving at time t enlarges property income at time $t + 1$ regardless of the value z_{t+1} takes (provided $k_t \in K$). A production function of the form $f(k, z) = z \cdot f^*(k)$ with $f^*(\cdot)$ Cobb–Douglas or CES with elasticity of substitution > 1 will satisfy (A5).

Assumptions (A1) to (A5) are sufficient to establish the existence of at least one invariant measure for the economy's equilibrium transition function v. The next assumption insures that the number of ergodic sets for each such v is finite and that none contain two or more cyclic subsets.

(A6) $z' > z$ implies $f(k, z') > f(k, z)$, $r(k, z') > r(k, z)$, and $w(k, z') > w(k, z)$ all $k > 0$; $z, z' \in Z$. $\tilde{z}$ has a continuous density function p, and $p(z) > 0$ all $z \in Z$.

Any production function of the form $f(k, z) = z \cdot f^*(k)$ with $f^*(\cdot)$ having conventional "neoclassical properties" will satisfy (A6).

---

[3] If $f(0, z) = 0$ all $z \in Z$, then $\{0\}$ is one ergodic set for our economy. We will not devote any attention to this particular ergodic set, however.

At the start of period t the state variable k_{t-1} summarizes the past history of the economy relevant for determining $w_t \cdot l_1(j)$ each j. Let $y\colon R^2 \to R^J$ be

$$y(k, z) = w(k, z) \cdot (l_1(1),\ldots, l_1(J)). \tag{4}$$

Provided young families at time t know z_t as they decide on $c_1(j)$ all j, k_t depends on $y_t = y(k_{t-1}, z_t)$ and expectations. We assume that all such families know y_t and use $g\colon R^J \to R$ with

$$g(y_t) = k_t \tag{5}$$

to predict k_t.[4] Then assuming that no family's net worth can ever become negative and that families attempt to maximize their expected utility, each family at time t chooses $c_1(j)$ to solve

$$\underset{\{c_1(j)\}}{\mathrm{Max}} \int_Z u(c_1(j), c_2(j), j)\, P(dz)$$

subject to:

$$\begin{aligned} c_1(j) + c_2(j)/r_{t+1} &= l_1(j) \cdot w_t + l_2(j) \cdot w_{t+1}/r_{t+1}, \\ r_{t+1} &= f_1(g(y_t), z), \\ w_{t+1} &= f(g(y_t), z) - f_1(g(y_t), z) \cdot g(y_t), \\ c_1(j) &\leqslant l_1(j) \cdot w_t. \end{aligned}$$

Suppose that

$$\gamma(y_t, g) = k_t \tag{6}$$

gives the actual value of k_t if all families base their expectations on g. Then the next section shows that for the sets $K \subset R$ and $Y \subset R^J$ defined below there is at least one $g\colon R^J \to R$ such that (i) $g(Y) \subset K$; (ii) $y(K \times Z) \subset Y$; and (iii) $g(y) = \gamma(y, g)$ all $y \in Y$. Thus for capital-to-labor ratios in K, the income vector for young families must lie in Y. For all income vectors in Y, however, anticipated next-period capital-to-labor ratios will again be in K and the anticipations will be borne out exactly.

To construct the sets Y and K we need two lemmas.

LEMMA 1. *There exists $k_U \in (0, \infty)$ such that $0 < k \leqslant k_U$ implies $\gamma(y(k, z), g) \leqslant k_U$ any $g\colon R^J \to R$ and all $z \in Z$.*

---

[4] Notice that k_t is not a stochastic variable here. Notice also that there is no benefit to one family of using observations of y_{t-s} any $s \geqslant 1$ in predicting k_t (unless it knows other families do).

LEMMA 2. *For any* $k_l < k_U$ *let* $K = K(k_L) = [k_L, k_U]$. *Let* $Y = \prod_{j=1}^{J} [y_L(j), y_U(j)]$, *where* $y_L(j) = w(k_L, z_L) \cdot l_1(j)$ *and* $y_U(j) = w(k_U, z_U) \cdot l_1(j)$ *all* j.[5] *Suppose* $y_L = (y_L(1),\dots, y_L(J))$ *and* $g\colon Y \to K$. *Then there exists* $k_L \in (0, k_U)$ *such that* $g(y_L) = k_L$ *implies* $\gamma(y_L, g) > k_L$.

The same $k_L \in (0, k_U)$ works for any $g\colon Y \to K$. We fix k_L, k_U, K and Y as they are defined in the lemmas.

II. EQUILIBRIUM

The operator γ (see line (6)) associates with each function $g\colon Y \to K$ a second function

$$g^* = \Gamma(g) \tag{7}$$

with $g^*(y) = \gamma(y, g)$ all $y \in Y$. Each fixed point for Γ defines an equilibrium (with rational expectations) for our model: if all families' expectations are given by g, the actual evolution of the economy is governed by $g^* = \Gamma(g)$. It is not difficult to prove that Γ must have at least one such fixed point.[6] In order to complete the analysis of steady states in Section III, however, we seek a fixed point $g = \Gamma(g)$ which is continuous and strictly monotone.

Our argument, therefore, requires several steps. Letting Y and K be as in Section I, and letting $\Delta = (\Delta_1,\dots, \Delta_J)$, define

$$\begin{aligned} G(\varepsilon) = \{ g\colon Y \to K \mid g &\text{ is continuous, } 0 \leqslant g(y + \Delta) - g(y) \leqslant \delta \cdot \varepsilon \\ &\text{all } y,\ y + \Delta \in Y \text{ with } \Delta_j = \delta \text{ some } j \text{ and } \Delta_i = 0 \text{ all } i \neq j \}. \end{aligned} \tag{8}$$

Then if $\|\cdot\|$ is the uniform norm on Y, $\|g - \bar{g}\| = \sup_{y\in Y} |g(y) - \bar{g}(y)|$.

LEMMA 3. *For any* $\varepsilon > 0$, $G(\varepsilon)$ *is compact in the topology induced by* $\|\cdot\|$.

Although Γ does not necessarily take $G(\varepsilon)$ into itself, we construct a continuous operator, Ψ, that does. For each $\varepsilon > 0$ the new function must have a fixed point on $G(\varepsilon)$. We show that with the correct choice of ε, however, all fixed points of Ψ on $G(\varepsilon)$ must also be fixed points of Γ.

---

[5] Although we are assuming that (A6) holds as we define $y_L(j)$ and $y_U(j)$, we could avoid the need to do that by defining $y_L(j) = w^*(k_L) \cdot l_1(j)$, where $w^*(k_L) = \text{Min}_{z \in Z}\, w(k_L, z)$, etc. The same applies later in the paper (until Section III) and in the Appendix.

[6] To see this, let $H = \{g\colon Y \to K\}$. Then H is compact in the topology of pointwise convergence and Γ is continuous on H with that topology. If $g \in H$ and $g^* = \Gamma(g)$, $g^*(y) < k_U$ all $y \in Y$ by Lemma 1. If $g(y) = k_L$, the proof of Lemma 2 can be modified to show $g^*(y) > k_L$ regardless of the choice of $y \in Y$. Thus, $\Gamma(\partial H) \subset H$. So, Γ has a fixed point on H.

First we establish three properties of any $g^* = \Gamma(g)$ with $g \in G(\varepsilon)$.

LEMMA 4. *Let* $g^* = \Gamma(g)$ *with* $g \in G(\varepsilon)$ *any* $\varepsilon > 0$. *Then* $g^*\colon R^J \to R$ *is continuous on* Y.

LEMMA 5. *Let* $g^* = \Gamma(g)$ *with* $g \in G(\varepsilon)$ *any* $\varepsilon > 0$. *Let* $\Delta = (\Delta_1, \dots, \Delta_J)$, $\Delta_j = \delta > 0$, $\Delta_i = 0$ *all* $i \neq j$, *and* $y, y + \Delta \in Y$. *Then* $g(y + \Delta) = g(y)$ *implies* $g^*(y + \Delta') > g^*(y)$ *each* $\Delta' \in (0, \Delta)$.

LEMMA 6. *Suppose* $g^* = \Gamma(g)$ *and* $g \in G(\varepsilon)$. *Let* $\Delta = (\Delta_1, \dots, \Delta_J)$ *with* $\Delta_j = \delta > 0$ *and* $\Delta_i = 0$ *all* $i \neq j$. *Suppose* $g(y) = g^*(y)$. *Let* y, $y + \Delta \in Y$. *Then there exists an* $\varepsilon^* > 0$ *such that* $\varepsilon = \varepsilon^*$ *and* $g(y + \Delta) - g(y) = \delta \cdot \varepsilon^*$ *some* $\delta > 0$ *imply* $g^*(y + \Delta') - g^*(y) < \delta' \cdot \varepsilon^*$ *some* $\Delta' \in (0, \Delta)$ *with* $\Delta'_j = \delta'$.

Lemma 4 shows that each g^* is continuous (with respect to the conventional Euclidean metric). Lemmas 5 and 6 will enable us to prove that the fixed points of Γ and Ψ are the same. From this point forward ε^* will refer to the constant derived in Lemma 6.

We need one more lemma before we can turn our attention to Ψ.

LEMMA 7. *Γ is continuous (with respect to $\|\cdot\|$) on $G(\varepsilon^*)$.*

We define Ψ as follows. Let y_L be as in Lemma 2, and let $g^* = \Gamma(g)$ with $g \in G(\varepsilon^*)$. Suppose $y(\delta) = (y_L(1) + \delta, y_L(2), \dots, y_L(J))$. For $y = y(\bar{\delta}) \in Y$ any $\bar{\delta} \geqslant 0$ define

$$h(y) = \operatorname*{Max}_{y_L \leqslant y(\delta) \leqslant y} \{ g^*(y(\delta)), k_L \}, \tag{9}$$

$$h^*(y) = \operatorname*{Min}_{0 \leqslant \delta \leqslant \bar{\delta}} \{ h(y(\delta)) + \varepsilon^* \cdot (\bar{\delta} - \delta) \}. \tag{10}$$

After selecting each $\bar{\delta}$ with $y \in Y$, choose any $y_1 \in | y_L(1), y_U(1) |$. Let $y(\delta) = (y_1, y_L(2) + \delta, y_L(3), \dots, y_L(J))$. For $y = y(\bar{\delta}) \in Y$ any $\bar{\delta} \geqslant 0$ define

$$h(y) = \operatorname*{Max}_{y_{L,1} \leqslant y(\delta) \leqslant y} \{ g^*(y(\delta)), h(y_{L,1}) \}, \tag{11}$$

where $y_{L,1} = (y_1, y_L(2), \dots, y_L(J))$. Define $h^*(y)$ using line (10). Repeat the step for each $y_1 \in | y_L(1), y_U(1) |$. Proceeding coordinate by coordinate, we have $h\colon Y \to R$ and $h^*\colon Y \to R$. Let

$$\psi(g) = h, \tag{12}$$

$$\Psi(g) = h^*. \tag{13}$$

Then

LEMMA 8. $\Psi(G(\varepsilon^*)) \subset G(\varepsilon^*)$.

We also have

LEMMA 9. *Ψ is continuous (with respect to $\|\cdot\|$) on $G(\varepsilon^*)$.*

Since Ψ is continuous, $\Psi(G(\varepsilon^*)) \subset G(\varepsilon^*)$, and $G(\varepsilon^*)$ is convex and compact, Schauder's fixed-point theorem (see Smart [20]) proves that there is at least one $g = \Psi(g) \in G(\varepsilon^*)$. The following proposition shows that $g = \Gamma(g)$ as well.

PROPOSITION I. *$g = \Psi(g) \in G(\varepsilon^*)$ implies $g = \Gamma(g)$. Thus, there is at least one $g \in G(\varepsilon^*)$ such that $g = \Gamma(g)$.*

If $g \in G(\varepsilon^*)$, then g is continuous and nondecreasing. In fact, Lemma 5 shows that $g = \Gamma(g) \in G(\varepsilon^*)$ implies g must be increasing (on Y).

If $g = \Gamma(g) \in G(\varepsilon^*)$ and if all families use g to predict the next-period aggregate capital stock, their behavior will bear their expectations out. So, if for each Borel set $A \subset R$,

$$v(k, A) = \Pr\{g(y(k, \tilde{z})) \in A\}, \tag{14}$$

then v will define a transition function for the economy consistent with rational expectations on the part of all consumers. The number $v(k, A)$ gives the probability that if $k_t = k$, k_{t+1} will be an element of the set A. The invariant measures for v are the topic of Section III.

Notice that our equilibrium expectation functions $g = \Gamma(g) \in G(\varepsilon^*)$ contrast to solutions of Arrow–Debreu–Bewley (see [1, 3, 5]) general equilibrium models if our consumers have different utility functions or labor endowments. For, suppose consumers are homogeneous. Then with a complete set of futures markets (for labor, investment securities, and output) all young workers at time t will save the same amount, buy even shares of output at time $t+1$ in all contingencies, agree to supply the same amount of labor in each contingency, and buy the same share of each investment security. Futures prices must adjust so that the capital stock, labor supply, and total sale of output will be the same at time $t+1$ regardless of the value z_{t+1} takes. In this situation there will be an Arrow–Debreu–Bewley equilibrium corresponding to each of our equilibria.

If consumers are heterogeneous, on the other hand, the equilibria of the two types of model will differ. With a complete set of futures markets, a highly risk averse family might at time t trade some of its next-period purchasing power contingent on the event of a favorable z_{t+1} to a less cautious family for extra purchasing power contingent on the event of an unfavorable z_{t+1}. In our model that cannot happen. Thus our equilibria will typically differ from those of a traditional general equilibrium model unless all families are identical.

III. STEADY STATES

If all families use $g = \Gamma(g) \in G(\varepsilon^*)$ to compute their expectations, v (defined in line (14)) gives the actual transition function for the aggregate capital stock. We are now ready to prove that v has a finite number of ergodic sets in K and that associated with each such ergodic set E_i there is a unique stationary distribution function D_i. Notice that if $K^{**} = K \cup (k_U, \infty)$, the proof of Lemma 1 shows $k_t \in K^{**}$ implies $k_{t+s} \in K$ some finite s with probability 1. The proof of Lemma 2 shows that any $k' \in (0, k_L)$ will work as a lower bound for K, with (k', k_L) being a collection of transitory states. Thus, if

$$K^* = (0, \infty), \tag{15}$$

the ergodic sets for v on K, $\{E_1, ..., E_N\}$, are the ergodic sets for v on K^*.

Since $g \in G(\varepsilon^*)$, K is an invariant set for v. Since $g \in G(\varepsilon^*)$ is continuous, Grandmont and Hildenbrand [8] establish the existence of at least one invariant measure for v on K, therefore. However, we will make the following argument to show that v has only a finite number of ergodic sets in K and that none decompose into two or more cyclic subsets.

We can easily establish that $v(k, \cdot)$ is a probability measure and $v(\cdot, A)$ is a measurable function.

LEMMA 10. *Let* $g = \Gamma(g) \in G(\varepsilon^*)$, *and let* v *be as in line* (14). *Then* $v(k, \cdot)$ *is a probability measure for each* $k \in K$, *and for each Borel* $A \subset K$ $v(\cdot, A)$ *is a measurable function on* K.

The paragraph surrounding Proposition I shows that $g = \Gamma(g) \in G(\varepsilon^*)$ implies g must be increasing as well as continuous. We can use that fact to prove that v has a finite number of ergodic sets in K.

PROPOSITION II. *Suppose* $g = \Gamma(g) \in G(\varepsilon^*)$ *and* v *is as defined in line* (14). *Then there are a finite number of ergodic sets* $E_1, ..., E_N \subset K$ *for* v. *Provided no* E_i *can be decomposed into two or more cyclic subsets, with each* E_i *we can associate a unique stationary distribution function* D_i.

The proof requires Lemma 10 and the theorems contained in Chapter 5 of Doob [6]. We will prove that no E_i can be decomposed into two or more cyclic subsets in a moment.

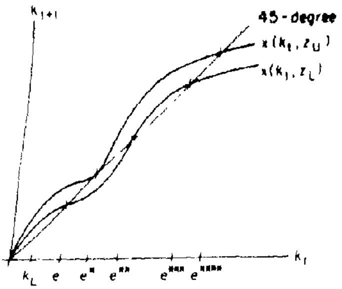


FIG. 1. The ergodic sets of v.

Define

$$x(k, z) = g(y(k, z)). \tag{16}$$

Then y is strictly increasing in both arguments, so the same is true for x. As Fig. 1 shows, Lemmas 1 and 2 imply that $x(k_L, z_L) > k_L$ and $x(k_U, z_U) \leqslant k_U$. Because of strict monotonicity, the x-curves for $z = z_L$ and z_U do not touch at any $k_t \in K$. In the diagram there are two ergodic sets: $E_1 = [e, e^*]$ and $E_2 = [e^{***}, e^{****}]$. Proposition II proves there must be at least one.

The points $k_t \in K - (E_1 \cup E_2)$ represent transitory states. If $k_t \in [k_L, e]$, then $\lim_{s\to\infty} \Pr\{k_{t+s} \in E_1\} = 1$. If $k_t \in [e^{**}, e^{***}] \cup [e^{****}, k_U]$, then $\lim_{s\to\infty} \Pr\{k_{t+s} \in E_2\} = 1$. If $k_t \in [e^*, e^{**}]$, then $k_{t+s} \in E_1$ or E_2 with probability 1 as $s \to \infty$. Thus, $k_t \in K$ (or K^*) implies the economy will end up in one ergodic set or the other with probability 1.

Once in an ergodic set such as E_1, k will oscillate systematically among the cyclic subsets of E_1 if two or more such subsets exist. In the latter case there will be one stationary distribution D_{1a} for each cyclic set. If there are n cyclic sets, $C_1, \ldots, C_n \subset E_1$, samplings at n-period intervals will fall in C_i with frequencies determined by D_{1i}. We rule out such a situation using assumption (A6).

PROPOSITION III. *Let $g = \Gamma(g) \in G(\varepsilon^*)$, and let v be as in line* (14). *Let $E \subset K$ be ergodic. Then E can have only one cyclic subset of positive measure.*

Thus, if $k_t \in K^*$, we know that with probability 1 k_{t+s} will enter an ergodic set $E \in \{E_1, \ldots, E_N\}$ with $E_i \subset K$ all i at some finite s. If $k_{t+s} \in E$, then $k_{t+s^*} \in E$ with probability 1 for each $s^* > s$. There will be a single distribution D associated with E such that $\int_A D(dk)$ gives the frequency with which realizations k_{t+s}, $s \geqslant 0$ will fall in any Borel set $A \subset E$ over the very long run.[7]

---

[7] Let $v^{(1)}(k, A) = v(k, A)$ and $v^{(i+1)}(k, A) = \int_K v^{(i)}(k^*, A)\, v(k, dk^*)$ all $i \geqslant 1$. Then $\int_A D(dk) = \lim_{i\to\infty} v^{(i)}(k, A)$ any Borel $A \subset E$ is a second way of looking at the role of D.

IV. Conclusions

We have proven the existence of an equilibrium transition function v for our economy consistent with rational expectations on the part of all households. We also have shown that with each such v we can associate N ergodic sets E_l (with $0 < N < \infty$) and accompanying stable, unique stationary distribution functions D_l.

Although the random shocks in our model are intertemporally independent, Lemma 5 shows that *for any equilibrium* $g \in G(\varepsilon^*)$ we cannot have $g_j^+(y) = 0$ any $y \in Y$ and j, so non-serially correlated sequences of outcomes $\{k_t, k_{t+1}, \ldots\}$ are impossible.[8] Thus, our model complements work done on "business-cycle models": models in the latter category generally study the role of shortsightedness or "nonrationality" (see, for example, Samuelson [19]) or incompletely revealed information (see, for example, Lucas [12] or [13]) in generating serially dependent outcome sequences from random demand shocks. Our shocks occur on the supply side of the model, but they generate serially dependent outcomes in spite of complete "rationality" in terms of behavior and expectations on the part of all market participants.

Appendix

1. *Proof of Lemma* 1. Assumption (A3) implies there exists $k_U \in (0, \infty)$ such that $\mathrm{Max}_{z \in Z}\{f(k_U, z)\} = k_U$. Let $k_t \in (0, k_U)$, and let $\iota = (1, \ldots, 1) \in R^J$. Then $k_{t+1} \leqslant \iota \cdot y(k_t, z_{t+1}) \leqslant f(k_t, z_{t+1}) \leqslant k_U$ any $z_{t+1} \in Z$. ∎

2. *Proof of Lemma* 2. Let "lim" stand for "$\lim_{k_L \downarrow 0}$." Suppose (A4) applies for $j = 1$. Let $g(y_L) = k_L$, and note that $y_L = y_L(k_L)$ depends on k_L. Let $s = y_L(1) - c_1$ if c_1 is the desired consumption now of young household $j = 1$ given y_L and g. Let $c_2(z)$ be the same household's anticipated consumption for next period if the next-period realization of $\tilde{z}$ is z.

Step 1. Let $c_U = w(k_L, z_U) \cdot l_2(1) + r(k_L, z_U) \cdot s$. Then $c_2(z) \leqslant c_U$ all $z \in Z$. (A1) and first-order conditions show $MRS(c_1, c_U) \equiv u_1(c_1, c_U, 1)/u_2(c_1, c_U, 1) \geqslant E[u_1(c_1, c_2(\tilde{z}), 1)]/E[u_2(c_1, c_2(\tilde{z}), 1)] \geqslant r(k_t, z_t)$ any k_L. $\lim r(k_L, r_L) = \infty$. So, $\lim MRS(c_1, c_U) = \infty$. Let $r = r(k_L, r_L)$, $r^* = r(k_L, z_U)$, $w = w(k_L, z_L)$, and $w^* = w(k_L, z_U)$. Note that $y_L(1) = w \cdot l_1(1)$.

Suppose $\lim w(k_L, z) = 0$ some $z \in Z$. Then (A4) implies the same holds for all $z \in Z$. $w \cdot l_1(1) \geqslant c_1 \geqslant 0$, so $\lim w = 0$ implies $\lim c_1 = 0$. Either $\lim c_U = 0$ or $\lim r^* s \neq 0$. In the latter case $\lim r^* s/w = \infty$. If $\lim c_U = 0$, $\lim MRS(c_1, c_U) = \infty$ and $\lim c_1 = \lim c_U = 0$. So, (A4) implies $\lim c_U/c_1 = \infty$.

---

[8] Recall that $g_j^+(y)$ stands for the partial derivative from the right for g at y.

$c_U/c_1 = [(w^*l_2(1)/w) + (r^*s/w)]/[l_1(1) - (s/w)]$. So, $\lim c_U/c_1 = \infty$ implies $\lim r^*s/w = \infty$ or $\lim s/w = l_1(1)$. The latter implies $\lim r^*s/w = \infty$. Thus, $\lim w(k_L, z) = 0$ all $z \in Z$ implies $\lim r^*s/w = \infty$.

Suppose $\lim w(k_l, z) > 0$ all $z \in Z$. Suppose also that $\lim r^*s/w \neq \infty$. Then $\lim s/w = 0 = \lim s$. So, $\lim c_1 = \lim wl_1(1) > 0$. For some sequence $(k_l)_n \to 0$ we must have $r^*s \to B \in [0, \infty)$ as $(k_l)_n \to 0$. Let "lim*" stand for "limit as $(k_l)_n \to 0$," so that $\lim^* r^*s = B$. Note that $\lim^* w^*l_2(1) = \lim w^*l_2(1) = A$ some $A \in [0, \infty)$. If $A + B > 0$, $\lim^* MRS(c_1, c_l) = MRS(\lim^* wl_1(1), A + B) < \infty$, a contradiction of $\lim MRS(c_1, c_l) = \lim^* MRS(c_1, c_l) = \infty$ established above. If $A + B = 0$, then $\lim^* c_1 > 0$, $\lim^* c_2 = 0$, and $\lim^* MRS(c_1, c_l) = \lim^* r = \infty$. That contradicts (A1). So, $\lim w(k_l, z) > 0$ all $z \in Z$ implies $\lim r^*s/w = \infty$.

Step 2. $\infty = \lim r^*s/w = l \cdot \lim(s/k_l) \cdot \lim(w^*/w) \cdot \lim \sigma(k_l, z_l)$. So, (A4) implies $\lim s/k_L = \infty$. Thus, $\lim \gamma(y_l(k_L), g)/k_l \geqslant \lim s/k_l = \infty$. So, there exists $k_L \in (0, k_l)$ with $\gamma(y_l(k_L), g) > k_l$ if $g(y_l(k_l)) = k_l$. ▮

3. *Proof of Lemma* 3. Since $G(\varepsilon)$ is closed (see Munkres [16, p. 130]), it is compact by Ascoli's theorem (see [16, p. 277]) if it is equicontinuous. It is "equicontinuous" if for any $\eta > 0$ and $y \in Y$ there exists $\delta > 0$ such that $g \in G(\varepsilon)$, $y' \in Y$, and $\mathrm{Max}_j |y_j - y'_j| < \delta$ implies $|g(y) - g(y')| < \eta$.

Fix any $\eta > 0$. Suppose $y, y' \in Y$. Define $y^* = (y^*_1, \dots, y^*_J)$ with $y^*_j = \mathrm{Min}\{y_j, y'_j\}$. Define y^{**} with $y^{**}_j = \mathrm{Max}\{y_j, y'_j\}$. Let $g \in G(\varepsilon)$. Then $g(y^*) \leqslant g(y)$, $g(y') \leqslant g(y^{**})$.

We have $y^*, y^{**} \in Y$. Suppose $\mathrm{Max}_j |y_j - y'_j| < \delta \equiv \eta/(J \cdot \varepsilon)$. Then $\mathrm{Max}_j |y^*_j - y^{**}_j| < \delta$. So, $|g(y) - g(y')| \leqslant g(y^{**}) - g(y^*) < \delta \cdot J \cdot \varepsilon = \eta$. Thus $G(\varepsilon)$ is equicontinuous. ▮

4. *Proof of Lemma* 4. Suppose $c_1(j) = c_1(j, y, g)$ maximizes $\xi(j, y, c_1, g) \equiv \int_Z u(c_1, \tilde{r} \cdot (y_j - c_1) + l_2(j) \cdot \tilde{w}, j)\, P(dz)$, where $\tilde{r} = r(g(y), \tilde{z})$, $w = w(g(y), \tilde{z})$, and $y = (y_1, \dots, y_J)$. Since ξ is concave in c_1 and continuous in y, Berge's [2] "maximum theorem" shows $c_1(j, y, g)$ is continuous in y each j. Thus $g^*(y) = \sum_{j=1}^{J} |y_j - c_1(j, y, g)|$ is continuous in y. ▮

5. *Proof of Lemma* 5. $g(y + \Delta) = g(y)$ implies the partial derivative from the right $g_j^+(y^*)$ exists and equals 0 all $y^* \in [y, y + \Delta)$. First-order conditions show $c_1(i, y + \Delta', g) = c_1(i, y, g)$ all $\Delta' \in (0, \Delta)$ and $i \neq j$.

Step 1. Let u_{11}, u_{12}, and u_{22} be the partial derivatives of $u(c_1(j), c_2(j), j)$ if $c_2(j) = \tilde{r} \cdot (y^*_j - c_1(j)) + \tilde{w} \cdot l_2(j)$, $\tilde{r} = r(g(y^*), \tilde{z})$, and $\tilde{w} = w(g(y^*), \tilde{z})$. Differentiating first-order conditions for $c_1(j)$, $E[u_{11} - 2 \cdot \tilde{r} \cdot u_{12} + \tilde{r}^2 \cdot u_{22}]\, dc_1(j) = E[-\tilde{r} \cdot u_{12} + \tilde{r}^2 \cdot u_{22}]\, dy_j$ at each $y^* \in [y, y + \Delta)$ if $dy_j > 0$. Then (A1) shows $dc_1(j)/dy_j < 1$. So, $g^{*+}_j(y^*) = 1 - dc_1(j)/dy_j > 0$.

Step 2. Thus, $g^{*+}_j(y^*) > 0$ each $y^* \in [y, y + \Delta)$. Since $g \in G(\varepsilon)$ and u is twice continuously differentiable, g^{*+}_j is continuous at each such y^*. So,

$g^*(y+\Delta') - g^*(y) \geqslant \mathrm{Min}_{y^* \in [y, y+\Delta']} g_j^{*+}(y^*) \cdot \delta' > 0$ any $\Delta' \in (0, \Delta)$ with $\Delta'_j = \delta'$. ▮

6. *Proof of Lemma* 6. Suppose $g(y+\Delta) - g(y) = \delta \cdot \varepsilon$ and $g \in G(\varepsilon)$. Then $g_j^+(y^*) = \varepsilon$ all $y^* \in [y, y+\Delta)$. Let $y = (y_1, ..., y_J)$ be the vector of incomes for young families now, and let $k = g(y)$. Let $s_m = y_m - c_1(m)$ all m, and let $s = \sum_{m=1}^J s_m$. Then $s_m \geqslant 0$ all m and $s = k$ if $g(y) = g^*(y)$. Let $g(y) = g^*(y)$.

Step 1. Let $\bar{r} = r(k_L, z_U)$. Assumption (A5) implies $r(k, z) + r_1(k, z) \cdot k > 0$ all $(k, z) \in K \times Z$. Thus there exists $r^* > \bar{r}$ such that $(r^*/(r^*+1)) \cdot r(k, z) + r_1(k, z) \cdot k > 0$ all $(k, z) \in K \times Z$. Fix $\varepsilon = 1 + r^*$.

Step 2. Set $dy_j = dx > 0$ and $dy_i = 0$ all $i \neq j$. Let $ds_j = dx + (r^* \cdot s_j/s)\, dx$ and $ds_i = (r^* \cdot s_i/s)\, dx$ all $i \neq j$. Then $ds \equiv \sum_{m=1}^J ds_m = \varepsilon \cdot dx$ and $ds_m \geqslant 0$ all m.

Step 3. Let $r = r(g(y), z)$, where z is the *next-period realization of* $\tilde{z}$. Then $d(r \cdot s_m)/dy_j = r(ds_m/dy_j) + s_m r_1 g_j^+(y) \geqslant (rr^* s_m/s) + s_m r_1 \varepsilon = \varepsilon \cdot |(r^*/\varepsilon) \cdot r + r_1 k| \cdot (s_m/s) \geqslant 0$ by Step 1 with strict inequality when $s_m > 0$.

Step 4. Step 3, (A1), and the first-order conditions for $c_1(m)$ show that if y changes to $y + \Delta'$, where $\Delta' \in (0, \Delta)$, and if each family's saving is changed by δ'_j, where $\delta'_j = \delta' + (r^* s_j/s)\, \delta'$ or $\delta'_i = (r^* s_i/s)\, \delta'$ when $i \neq j$, then provided $\delta' = \Delta'_j$ is small enough, each family of type m will want to decrease the change in its saving if $s_m > 0$ or $m = j$ and leave it unchanged otherwise.

Suppose δ_m^* is the desired change in saving for family m following the change from y to $y + \Delta'$ above. Suppose $\sum_{m=1}^J \delta_m^* > \sum_{m=1}^J \delta'_m$. Let $k^* = k + \sum_{m=1}^J \delta_m^*$ and $k' = k + \sum_{m=1}^J \delta'_m$. Then $k^* > k'$ so (A5) implies $r(k^*, z) \cdot k^* > r(k', z) \cdot k'$ all $z \in Z$ (assuming $k^*, k' \in K$). Thus for some m we have $r(k^*, z) \cdot (s_m + \delta_m^*) > r(k', z) \cdot (s_m + \delta'_m)$ any $z \in Z$. So, the arguments above show the first-order conditions of family m are being violated at $s_m + \delta_m^*$, a contradiction. So, $\sum_{m=1}^J \delta_m^* > \sum_{m=1}^J \delta'_m$ is not possible.

If $0 < \sum_{m=1}^J \delta_m^* \leqslant \sum_{m=1}^J \delta'_m$, the first paragraph of Step 4 shows $\delta_m^* \leqslant \delta'_m$ all m, with strict inequality for at least one m. Thus, $g^*(y+\Delta') - g^*(y) = \sum_{m=1}^J \delta_m^* < \sum_{m=1}^J \delta'_m = \varepsilon \cdot \delta'$. ▮

7. *Proof of Lemma* 7. Suppose $g, \bar{g} \in G(\varepsilon^*)$ and $\| g - \bar{g} \| = \delta$. Let $g^* = \Gamma(g)$ and $\bar{g}^* = \Gamma(\bar{g})$.

Maximizing over $(k, z) \in K \times Z$, let $r_1^* = \mathrm{Max}\{|r_1(k, z)|\}$ and $w_1^* = \mathrm{Max}\{w_1(k, z)\}$. Let s_i be the saving of young household i at present given g, and let $\bar{s}_i$ be the same given $\bar{g}$. Then first-order conditions show $|s_i - \bar{s}_i| \leqslant \delta x$ for all i where $x = k_U r_1^* + l w_1^*$. So, $\| g^* - \bar{g}^* \| \leqslant \delta \cdot x \cdot J$. Thus, Γ is continuous. ▮

8. *Proof of Lemma* 8. Let $y, y' \in Y$. Set $y_i^* = \text{Min}\{y_i, y_i'\}$ and $y_i^{**} = \text{Max}\{y_i, y_i'\}$ all i. Then $y^* \leqslant y, y' \leqslant y^{**}$ and $y, y^* \in Y$. We have $|h^*(y) - h^*(y')| \leqslant h^*(y^{**}) - h^*(y^*) \leqslant \sum_{i=1}^{J} [h^*(y^{**}(i)) - h^*(y^*(i))]$, where $y^*(1) = y^*$, $y^*(2) = (y_1^{**}, y_2^*, \ldots, y_J^*)$, $y^*(3) = (y_1^{**}, y_2^{**}, y_3^*, \ldots, y_J^*), \ldots$, $y^{**}(1) = y^*(2)$, $y^{**}(2) = y^*(3), \ldots$, and $y^{**}(J) = y^{**}$. Lines (9)–(11) and Lemma 4 show the terms in brackets are continuous. $\blacksquare$

9. *Proof of Lemma* 9. Let $g, \bar{g} \in G(\varepsilon^*)$. Let $h = \psi(g)$, $h^* = \Psi(g)$, $\bar{h} = \psi(\bar{g})$, $\bar{h}^* = \Psi(\bar{g})$, $g^* = \Gamma(g)$, and $\bar{g}^* = \Gamma(\bar{g})$.

Step 1. Claim $\| g^* - \bar{g}^* \| \geqslant \| h - \bar{h} \|$. Choose any $y \in Y$. Suppose $h(y) > \bar{h}(y)$. Then $h(y) = g^*(x)$ some $x \leqslant y$, $x \in Y$ or $h(y) = k_L$. In the latter case $h(y) = \bar{h}(y)$. Otherwise, $|h(y) - \bar{h}(y)| = h(y) - \bar{h}(y) \leqslant g^*(x) - \bar{g}^*(x) \leqslant \| g^* - \bar{g}^* \|$. Since this holds for any $y \in Y$, $\| h - \bar{h} \| \leqslant \| g^* - \bar{g}^* \|$.

Step 2. Claim $\| h^* - \bar{h}^* \| \leqslant \| h - \bar{h} \|$. Let $y \in Y$. Suppose $h^*(y) > \bar{h}^*(y)$. Then $\bar{h}^*(y) = \bar{h}(x) + \varepsilon^* \cdot \sum_{j=1}^{J} (y_j - x_j)$ some x with $y_L \leqslant x \leqslant y$. So, $|h^*(y) - \bar{h}^*(y)| = h^*(y) - \bar{h}^*(y) \leqslant h(x) + \varepsilon^* \cdot \sum_{j=1}^{J} (y_j - x_j) - \bar{h}(x) - \varepsilon^* \cdot \sum_{j=1}^{J} (y_j - x_j) = h(x) - \bar{h}(x) \leqslant \| h - \bar{h} \|$.

Step 3. Let any $\eta > 0$ be given. Lemma 7 implies there exists a $\delta > 0$ such that $\| g^* - \bar{g}^* \| < \eta$ if $\| g - \bar{g} \| < \delta$. Combining this with steps 1 and 2 above, $\| g - \bar{g} \| < \delta$ implies $\| h^* - \bar{h}^* \| < \eta$. Thus Ψ is continuous. $\blacksquare$

10. *Proof of Proposition* I. Let $g = \Psi(g) \in G(\varepsilon^*)$. Let $g^* = \Gamma(g)$.

Step 1. $g^*(y_L) = g(y_L)$ unless $g^*(y_L) < k_L = g(y_L)$. But, Lemma 2 rules out the latter. Hence, $g^*(y_L) = g(y_L)$, and in fact both exceed k_L.

Step 2. Let $x(\delta) = (y_L(1) + \delta, y_L(2), \ldots, y_L(J))$. Let $\bar{\delta} = \inf\{\delta > 0 \mid g^*(x(\delta)) \neq g(x(\delta))\}$.

Suppose $g^*(x(\bar{\delta} + \eta)) < g^*(x(\bar{\delta}))$ all $\eta \in (0, \bar{\eta})$ some $\bar{\eta} > 0$. Then $g = \Psi(g)$ implies $g(x(\bar{\delta} + \eta)) = g(x(\bar{\delta}))$ all such η. So, Lemma 5 implies $g^*(x(\bar{\delta} + \eta)) > g^*(x(\bar{\delta}))$ some $\eta \in (0, \bar{\eta})$, a contradiction.

Suppose $g^*(x(\bar{\delta} + \eta)) - g^*(x(\bar{\delta})) > \varepsilon^* \cdot \eta$ all $\eta \in (0, \bar{\eta})$ some $\bar{\eta} > 0$. Then $g = \Psi(g)$ implies $g(x(\bar{\delta} + \eta)) - g(x(\bar{\delta})) = \varepsilon^* \cdot \eta$ all $\eta \in (0, \bar{\eta})$. Then Lemma 6 implies $g^*(x(\bar{\delta} + \eta)) - g^*(x(\bar{\delta})) < \varepsilon^* \cdot \eta$ some $\eta \in (0, \bar{\eta})$, a contradiction of our supposition.

So, $g(x(\delta)) = g^*(x(\delta))$ all $\delta \geqslant 0$.

Step 3. Step 2 applies for $x(\delta) = (y_1, y_L(2) + \delta, y_L(3), \ldots, y_L(J))$ any $y_1 \in [y_L(1), y_U(1)]$. And, continuing coordinate by coordinate, for $x(\delta) = (y_1, \cdot, y_{i-1}, y_L(i) + \delta, y_L(i+1), \ldots, y_L(J))$ any $i \leqslant J$ and any $(y_1, \ldots, y_{i-1}) \in \prod_{j=1}^{i-1} [y_L(j), y_U(j)]$. Thus, $g(x) = g^*(x)$ all $x \in Y$. $\blacksquare$

11. *Proof of Lemma* 10. $v(k, \cdot)$ is a probability measure by definition. Suppose $A \subset K$ is a closed interval. Lemma 5 shows $g = \Gamma(g)$ implies g is

increasing. So, $B(k) = \{z \in Z \mid g(y(k, z)) \in A\}$ is a bounded, closed interval. Because g and y are continuous, B is upper semi-continuous. Thus, $v(k, A) = \int_{B(k)} P(dz)$ is continuous in k by (A6). So, $v(\cdot, A)$ is measurable.

The class of closed intervals generates the class of Borel sets of K. Thus, for any Borel $A \subset K$, any $k \in K$, and any positive integer n, we can find disjoint, closed intervals $C_1, \ldots, C_{m(n)}$ with $A \subset \bigcup_{i=1}^{m(n)} C_i$ and $0 \leqslant v(k, \bigcup_{i=1}^{m(n)} C_i) - v(k, A) < 1/n$. Yet, $v(k, \bigcup_{i=1}^{m(n)} C_i) = \sum_{i=1}^{m(n)} v(k, C_i)$ is measurable by the preceding argument. So we have a sequence of measurable functions $v(k, \bigcup_{i=1}^{m(n)} C_i) \to v(k, A)$ pointwise for each $k \in K$. Thus, $v(\cdot, A)$ is measurable on K (see Halmos [10, p. 84]). ▮

12. *Proof of Proposition* II. Let μ be the Lebesgue measure. Suppose there exists a constant $\eta > 0$ such that for any Borel $A \subset K$ with $\mu(A) < \eta$ we have $v(k, A) < 1 - \eta$ all $k \in K$. Then Chapter 5 of Doob [6] shows that given Lemma 10 we will have N ergodic sets with $N < \mu(K)/\eta$. Doob's Theorem 5.7 shows that with each E_i we can associate a unique stationary D_i given our hypotheses.

We must show that $\eta > 0$ exists.

Step 1. Let $x(k, z) = g(y(k, z))$ and $\phi(k, z', z) = x(k, z') - x(k, z)$. Then ϕ is continuous and $\phi(k, z', z) > 0$ if $k \in K$; $z, z' \in Z$; and, $z' > z$. Let $\phi^* = \mathrm{Min}_{k \in K}\{\phi(k, z_U, z_L)\}$. Then $\phi^* > 0$.

Step 2. Note that $\mu(K) \geqslant \phi^*$. Let $\zeta = \phi^*/2$. Suppose $A \subset K$ and $\mu(A) < \zeta$. Let $B(k) = \{z \in Z \mid x(k, z) \in A\}$ and $C(k) = Z - B(k)$. Let $\psi(k, z, \delta) = \phi(k, z + \delta, z)/\delta$. Let $W^* = \mathrm{Max}_{(k,z) \in K \times Z}\{l \cdot w_2(k, z)\}$. Then $\psi(k, z, \delta) \leqslant \varepsilon^* W^*$ all $k \in K$; $z, z + \delta \in Z$; and, $\delta > 0$. So, $\mu(C(k)) \cdot \varepsilon^* \cdot W^* \geqslant \mu(K - A) = \mu(K) - \mu(A) > \phi^* - \zeta = \zeta$. Thus, $\mu(C(k)) > \zeta/(\varepsilon^* W^*)$.

Assumption (A6) implies $p(z) \geqslant p_L$ all $z \in Z$ some $p_L > 0$. Thus, $v(k, A) = \int_{B(k)} p(z)\, dz = \int_Z p(z)\, dz - \int_{C(k)} p(z)\, dz < 1 - p_L \zeta/(\varepsilon^* W^*)$ any $k \in K$.

Step 3. Let $\eta = \mathrm{Min}\{\zeta, p_L \zeta/(\varepsilon^* W^*)\}$. Then $\mu(A) < \eta$ and $A \subset K$ imply $v(k, A) < 1 - \eta$ any $k \in K$. ▮

13. *Proof of Proposition* III. Suppose C is a cyclic subset of E and μ is the Lebesgue measure. Let $k \in E$ with $v(k, C) = 1$. Figure 1 shows $x(k, z) = g(y(k, z))$ implies $x(k, z_U) \geqslant k \geqslant x(k, z_L)$. Then $I = (x(k, z_L), x(k, z_U))$ implies $\mu(I) > 0$ (by Lemma 5) and $\mu(I \cap C) = \mu(I)$.

Let $k^* \in I \cap C$. $x(k^*, z_U) \geqslant k^* \geqslant x(k^*, z_L)$. Let $B(k^*) = \{z \in Z \mid x(k^*, z) \in I \cap C\}$. Then $\mu(B(k^*)) > 0$. So, $v(k^*, C) \geqslant v(k^*, I \cap C) = \int_{B(k^*)} p(z)\, dz > 0$ by (A6). Since this result holds for any $k^* \in I \cap C$, and since $\mu(I \cap C) > 0$, there cannot be a second cyclic subset of positive measure.[9] ▮

---

[9] In other words, we have shown that the period of C is 1.

ACKNOWLEDGMENT

I wish to thank my colleague Larry Blume for helpful comments at an early stage in this research.

REFERENCES

1. K. J. ARROW AND F. H. HAHN, "General Competitive Analysis," Holden–Day, San Francisco, 1971.
2. C. BERGE, "Topological Spaces," trans. E.M. Patterson, Oliver & Boyd, Edinburgh, 1963.
3. T. F. BEWLEY, Existence of equilibria in economies with infinitely many commodities, *J. Econ. Theory* **4** (1972), 514–540.
4. W. BROCK AND L. MIRMAN, Optimal economic growth and uncertainty: the discounted case, *J. Econ. Theory* **4** (1972), 479–513.
5. G. DEBREU, "Theory of Value," Yale Univ. Press, New Haven, 1973.
6. J. DOOB, "Stochastic Processes," Wiley, New York, 1953.
7. J.M. GRANDMONT, Temporary general equilibrium theory, *Econometrica* **45** (1977), 535–572.
8. J.M. GRANDMONT AND W. HILDEBRAND, Stochastic processes of temporary equilibria, *J. Math. Econ.* **1** (1974), 247–278.
9. J. R. GREEN AND M. MAJUMDAR, The nature of stochastic equilibria, *Econometrica* **43** (1975), 647–660.
10. P. R. HALMOS, "Measure Theory," Springer–Verlag, New York, 1974.
11. F. S. INABA, Consumption loans under uncertainty, *J. Econ. Theory* **15** (1977), 26–53.
12. R. E. LUCAS JR., Expectations and the neutrality of money, *J. Econ. Theory* **4** (1972), 103–124.
13. R. E. LUCAS JR., An equilibrium model of the business cycle, *J. Pol. Econ.* **83** (1975), 1113–1144.
14. L. J.MIRMAN, On the existence of steady state measures for one sector growth models with uncertain technology, *Internat. Econ. Rev.* **13** (1972), 271–286.
15. L. J. MIRMAN, The Steady State behavior of a class of one sector growth models with uncertain technology, *J. Econ. Theory* **6** (1973), 219–242.
16. J. R. MUNKRES, "Topology," Prentice–Hall, Englewood Cliffs, N. J., 1975.
17. R. RADNER, Existence of equilibrium of plans, prices, and price expectations in a sequence of markets, *Econometrica* **40** (1972), 289–304.
18. R. RADNER, Market equilibrium and uncertainty: concepts and problems, *in* "Frontiers of Quantitative Economics, Vol. II" (M. D. Intriligator and D. A. Kendrick, Eds.), North–Holland, Amsterdam, 1974.
19. P. A. SAMUELSON, Interactions between the multiplier analysis and the principle of acceleration, *Rev. Econ. Statist.* **21** (1939), 75–78.
20. D. R. SMART, "Fixed Point Theorems," Cambridge Univ. Press, London, 1974.

JOURNAL OF ECONOMIC THEORY 24, 393–397 (1981)

Existence of Budget Constrained Pareto Efficient Allocations

HANS KEIDING

Institute of Economics,
University of Copenhagen, Copenhagen DK-1455, Denmark

Received February 26, 1980; revised June 2, 1980

The concept of a budget constrained Pareto efficient (BCPE) allocation was introduced by Balasko [1], who gave an existence proof in the differentiable case using a duality approach. Intuitively, a BCPE allocation $(x_1,..., x_m)$ is Pareto efficient, and, in addition, the bundle of consumer i satisfies a budget constraint $qx_i = q\omega_i$, where $q \in \mathbb{R}^l_+$ is a given, fixed price vector and ω_i the initial endowment of consumer i.

In the present note we give a direct proof of the existence of BCPE allocations under rather weak conditions: thus preferences are not necessarily total or transitive. Only convexity of the set of preferred bundles together with some monotonicity and continuity conditions are assumed, and we use the approach to existence proofs initiated by Gale and Mas–Colell [3] and Shafer and Sonnenschein [4].

An application of the model to the problem of fair division of a random harvest introduced by Gale [2] is indicated.

Let $\mathscr{E} = ((X_i, P_i)_{i=1}^m, \omega)$ be an exchange economy, where for $i = 1,..., m$:

$X_i = \mathbb{R}^l_+$ is the consumption set; let $X = \prod_{i=1}^m X_i$;
$P_i: X \to X_i$ is the preference correspondence,

assigning to each allocation $x = (x_1,..., x_m) \in X$ a non-empty subset $P_i(x)$ of X_i interpreted as the bundles preferred by consumer i to x_i given the bundles x_j of the other consumers. We assume that

(i) $\forall x \in X$, $x_i \notin P_i(x)$ (irreflexivity).
(ii) $\forall x \in X$, $P_i(x)$ is convex, and $x_i + (\mathbb{R}^l_+ \setminus \{0\}) \subset P_i(x)$.
(iii) Graph $P_i = \{(x, x_i') \in X \times X_i \mid x_i' \in P_i(x)\}$ is open in $X \times X_i$.

Condition (ii) says that the set of preferred bundles is convex, and that more of any commodity is preferred (monotonicity).

Finally, $\omega \in \mathbb{R}^l$ is the initial endowment available in the economy $\mathscr{E}$. We assume that $\omega \in \text{int}\, \mathbb{R}^l_+$.

0022-0531/81/030393–05$02.00/0
Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved

For each natural number n, we put

$$\Delta^{n-1} = \left\{ x \in \mathbb{R}^n_+ \mid \sum_{h=1}^{n} x_h = 1 \right\}.$$

Let $\theta = (\theta_1, \ldots, \theta_m) \in \Delta^{m-1}$, $q = (q_1, \ldots, q_l) \in \Delta^{l-1}$. We say that $(x, p) \in X \times \Delta^{l-1}$ is a *(θ, q)-constrained market equilibrium* if

(1) $\sum_{i=1}^{m} x_i = \omega$,

(2) $\forall i\colon qx_i = \theta_i q\omega$,

(3) $\forall i\colon x_i' \in P_i(x)$ implies $px_i' > px_i$.

THEOREM. *Let $\theta \in \operatorname{int} \Delta^{m-1}$, $q \in \operatorname{int} \Delta^{l-1}$. Then there exists a (θ, q)-constrained market equilibrium in $\mathscr{E}$.*

Proof. Choose $K > 0$ such that for $x \in \mathbb{R}^l_+$, $x_h = K$ implies $qx > q\omega$, all h. Let $\tilde{X}_i = X_i \cap \{x \in \mathbb{R}^l_+ \mid x_h \leqslant K,\ h = 1, \ldots, l\}$ and $\tilde{X} = \prod_{i=1}^{m} \tilde{X}_i$.

For $i = 1, \ldots, m$, define $r_i\colon \tilde{X} \to \mathbb{R}$ by

$$r_i(x) = \theta_i q\omega - qx_i$$

and $w_i(x, p) = \max\{\tfrac{1}{2}px_i,\ px_i + r_i(x)\}$. Then $w_i(p, x)$ is continuous, so the correspondence $\gamma_i\colon \tilde{X} \times \Delta \to \tilde{X}_i$ defined by

$$\gamma_i(x, p) = \{x_i' \in \tilde{X}_i \mid px_i' < w_i(x, p)\}$$

has open graph and convex (possibly empty) values. Define the correspondence $\phi_i\colon \tilde{X} \times \Delta^{l-1} \to \tilde{X}_i$ by

$$\begin{aligned} \phi_i(x, p) &= \gamma_i(x, p) \cap P_i(x) && \text{if } x_i \in \operatorname{cl} \gamma_i(x, p), \\ &= \gamma_i(x, p) && \text{otherwise.} \end{aligned}$$

Then ϕ_i has open graph and convex values, and $x_i \notin \phi_i(x, p)$, all $(x, p) \in \tilde{X} \times \Delta^{l-1}$.

Finally, let $\phi_0\colon \tilde{X} \times \Delta^{l-1} \to \Delta^{l-1}$ be the correspondence defined by

$$\phi_0(x, p) = \left\{ p' \in \Delta^{l-1} \mid p'\left(\sum_{i=1}^{m} x_i - \omega\right) > \max\left\{ p\left(\sum_{i=1}^{m} x_i - \omega\right), 0 \right\} \right\}.$$

Then ϕ_0 has open graph and convex values, and $p \notin \phi_0(x, p)$.

By the fixed point theorem of Gale and Mas-Colell [3, p. 10] there is $(x, p) \in \tilde{X} \times \Delta^{l-1}$ such that $\phi_i(x, p) = \varnothing$, $i = 1, \ldots, m$.

For all $i = 1,\dots, m$, $r_i(x) \leqslant 0$, since $r_i(x) > 0$, some i, implies $w_i(x, p) > px_i$, whence $\gamma_i(x, p) \neq \varnothing$. Since $x_{ih} = K$, some h, implies $qx_i > q\omega > \theta_i q\omega$ contradicting $r_i(x) > 0$, we must have $x_{ih} < K$, all h, whence $x_i + (\mathbb{R}^l_+ \setminus \{0\}) \cap \gamma_i(x, p) \neq \varnothing$ and $\phi_i(x, p) \neq \varnothing$. Thus $r_i(x) \leqslant 0$, all i.

Suppose now, that $\gamma_i(x, p) = \varnothing$, all i. Then $px_i = 0$, $i = 1,\dots, m$, and $p\sum_{i=1}^m x_i = 0$, so we cannot have $\sum_{i=1}^m x_i > \omega$. If there is some i such that $\gamma_i(x, p) \neq \varnothing$, then $r_i(x) = 0$, so $x_{ih} < K$, $h = 1,\dots, l$, and then $\gamma_i(x, p) \cap P_i(x) = \varnothing$ implies that $p \in \operatorname{int} \Delta^{l-1}$. For each i, $x_i \neq 0$, since $qx_i \geqslant \theta_i q\omega \neq 0$. But then $px_i > 0$, $i = 1,\dots, m$ whence $r_i(x) = 0$, all i, or $q\sum_{i=1}^m x_i = q\omega$. So, again we cannot have $\Sigma_{i=1}^m x_i > \omega$.

From $\phi_0(x, p) = \varnothing$, we get that either $\sum_{i=1}^m x_i \leqslant \omega$ or p maximizes $p'(\sum_{i=1}^m x_i - \omega)$ on Δ^{l-1}. If $\sum_{i=1}^m x_i \nleqslant \omega$, since $\sum_{i=1}^m x_i - \omega \notin \operatorname{int} \mathbb{R}^l_+$, it must have some non-positive coordinates so the maximizing p belongs to $bd\Delta^{l-1}$. If now $\gamma_i(x, p) = \varnothing$, all i, $px_i = 0$, so $p_h > 0$ implies $x_{ih} = 0$, $i = 1,\dots, m$. Then $\sum_{i=1}^m x_{ih} - \omega_h < 0$, and if p maximizes $p'(\sum_{i=1}^m x_i - \omega)$, $p_h = 0$, a contradiction. If there is i with $\gamma_i(x, p) \neq \varnothing$, we saw that $p \in \operatorname{int} \Delta^{l-1}$. So, we must have $\sum_{i=1}^m x_i \leqslant \omega$. Since $r_i(x) \leqslant 0$, all i, we must have $\sum_{i=1}^m x_i = \omega$.

From this we get that $px_i > 0$ for at least one i, whence $\gamma_i(x, p) \neq \varnothing$ and consequently $p \in \operatorname{int} \Delta^{l-1}$. Then $px_i > 0$, all i, each $\gamma_i(x, p)$ is non-empty, and $\phi_i(x, p) = \varnothing$ implies that $r_i(x) = 0$, i.e., $qx_i = \theta_i q\omega$, each i, and that $\gamma_i(p, x) \cap P_i(x) = \varnothing$. Since $P_i(x)$ is open in X_i, $\operatorname{cl} \gamma_i(p, x) \cap P_i(x) = \varnothing$, so $x'_i \in P_i(x)$ implies $px'_i > px_i$. Q.E.D.

Let $\theta \in \Delta^{m-1}$, $q \in \Delta^{l-1}$. We say that $x \in X$ is a (θ, q)-constrained Pareto efficient allocation if

(i) $\sum_{i=1}^m x_i = \omega$,

(ii) $\forall i$: $qx_i = \theta_i q\omega$,

(iii) x is Pareto efficient, i.e., there is no $x' \in X$ such that $\sum_{i=1}^m x'_i = \omega$ *and* $x'_i \in P_i(x)$, $i = 1,\dots, m$.

Corollary 1. *Let* $\theta \in \Delta^{m-1}$, $q \in \Delta^{l-1}$. *Then there exists a* (θ, q)-*constrained Pareto-efficient allocation.*

Proof. If $\theta \in \operatorname{int} \Delta^{m-1}$, $q \in \operatorname{int} \Delta^{l-1}$, there is a (θ, q)-constrained market equilibrium (x, p). Clearly, x satisfies (i) and (ii), and for any $x' \in X$ with $x'_i \in P_i(x)$, $i = 1,\dots, m$, we have $px'_i > px_i$, so $p(\sum_{i=1}^p x'_i - \omega) > p(\sum_{i=1}^m x_i - \omega)$ or $\sum_{i=1}^m x'_i \neq \omega$, so x is Pareto-efficient.

If $\theta \in bd\Delta^{m-1}$, $q \in \operatorname{int} \Delta^{l-1}$, let $I = \{i \mid \theta_i \neq 0\}$, put $x_i = 0$ for $i \notin I$ and repeat the above argument for the economy $\mathscr{E}' = ((X_i, P_i)_{i \in I}, \omega)$.

If $\theta \in \Delta^{m-1}$, $q \in bd\Delta^{l-1}$, let (q^n) be a sequence with $q^n \in \operatorname{int} \Delta^{l-1}$ converging to q. For each n, let x^n be a (θ, q^n)-constrained Pareto-efficient allocation. Since each x^n is in the compact set $H=\{x\in X \mid \sum_{i=1}^m x_i=\omega\}$, there is a subsequence of (x^n) converging to some $x^*\in H$. For each i, $qx_i^*=\lim q^n x_i^n=\lim \theta_i q^n\omega=\theta_i q\omega$, and since the set of Pareto-efficient allocations $\{x\in X \mid x\in H \text{ and } P_1(x)\times\cdots\times P_m(x)\cap H=\Phi\}$ is closed, x^* is Pareto-efficient. Q.E.D.

Remark 1. For $\theta=(\theta_1,\ldots,\theta_m)\in\Delta^{m-1}$, let (x,p) be an ordinary competitive equilibrium in the private ownership economy $\mathscr{E}^{\mathrm{pr}}=(X_i,P_i,\omega_i)_{i=1}^m$, where $\omega_i=\theta_i\omega$, $i=1,\ldots,m$. Then x is a (θ,p)-constrained Pareto-efficient allocation, and if preferences are complete, it is the only one. Balasko [1] shows that this uniqueness holds in a neighborhood of (θ,p). None of this results carry over to the general case of non-ordered preferences.

For another application of the theorem, consider the problem of dividing a random harvest: A resource such as an orchard is owned jointly by m agents, each agent having a share θ_i, $i=1,\ldots,m$. There are l states of nature, the probability of state h being q_h, $h=1,\ldots,l$, and the yield of the resource in state h is ω_h. A distribution scheme is a family $x=(x_{ih})_{i=1}^m{}_{h=1}^l$, where x_{ih} is the amount of the yield in state h allocated to agent i. It is said to be feasible if $\sum_{i=1}^m x_{ih}=\omega_h$ for each h.

Assume that each agent i has preferences on the random prospects $(x_{i1},\ldots,x_{il})$, i.e., to every distribution scheme x there is a set $P_i(x)$ of random prospects preferred to $(x_{i1},\ldots,x_{il})$ by agent i.

Following Gale [2], we say that a feasible distribution x is *fair* if

(1) x is Pareto-efficient, i.e., there is no feasible distribution scheme x' such that $(x'_{i1},\ldots,x'_{il})\in P_i(x)$, $i=1,\ldots,m$,

(2) for each agent i, $\sum_{h=1}^l q_h x_{ih}=\theta_i\sum_{h=1}^l q_h\omega_h$, i.e., the mean value of the amounts delivered to agent i equals his share of the mean value of the yield.

Define the contingent commodity h ($h=1,\ldots,l$) as "Yield of the resource if state h obtains." Then the random prospects of agent i can be viewed as (contingent) commodity bundles $x_i=(x_{i1},\ldots,x_{il})$, distribution schemes become allocations, and the total yield of the resource in the various states of nature is the endowment vector $\omega=(\omega_1,\ldots,\omega_l)$ of contingent commodities. Thus the random harvest model is formally equivalent to an exchange economy $\mathscr{E}$ as defined above. Moreover, a fair distribution scheme is a (θ,q)-constrained Pareto-efficient allocation in $\mathscr{E}$. Applying Corollary 1, we get

COROLLARY 2. *There exists a fair distribution scheme.*

Remark 2. In his original model [2], Gale used stronger assumptions on preferences, which implied uniqueness as well as existence of fair distribution schemes. As noted above, in the present context uniqueness does not necessarily obtain.

REFERENCES

1. Y. BALASKO, Budget constrained Pareto Efficient allocations, *J. Econ. Theory* **21** (1979), 359–379.
2. D. GALE, "Fair Division of a Random Harvest," Discussion Paper, 1977.
3. D. GALE AND A. MAS-COLELL, An equilibrium existence theorem for a general model without ordered preferences, *J. Math. Econ.* **2** (1975), 9–15.
4. W. SHAFER AND H. SONNENSCHEIN, Some theorems on the existence of competitive equilibrium, *J. Econ. Theory* **11** (1975), 83–93.

JOURNAL OF ECONOMIC THEORY 24, 398–412 (1981)

Prudence versus Sophistication in Voting Strategy

HERVÉ MOULIN

Université Paris 9, Centre de Recherches de Mathématique de la Décision, 75775 Paris Cedex 16, France

Received April 11, 1978; revised January 31, 1980

INTRODUCTION

Since voting schemes do induce strategic manipulation (by the Gibbard–Satterthwaite result) a natural but obscure question is What kind of manipulation? Answering this question amounts to describing the strategical behaviour of the agents, that is, to select an equilibrium concept for games in strategical form. Most of the literature devoted to this problem (see [1, 5, 6, 8, 11]) has dealt with non-cooperative behaviour of the agents. This is because the major concern has been the implementation problem which is stated as follows: Given any social choice function (or correspondence), that is, a particular collective decision making, is it possible to distribute privately the decision power among the agents in such a way that by exercising (non-cooperatively) this power, the agents eventually select the very outcome(s) recommended by the social choice function (correspondence). Any answer to this problem throws some light on the collective implications of non-cooperative behaviour, a question underlying most of the economic literature on collective decision mechanisms (see [4, 3] for a survey of the literature).

In this paper we assume that the agents behave non-cooperatively. This assumption alone is not enough to determine unambiguously the outcome elected by the agents. Given a particular voting scheme and a particular preference profile of the agents, there are in general several Nash equilibriums in the corresponding game and accordingly several different outcomes are the possible result of an equilibrium vote of the agents.[1] In order to make the behaviour of the agents well defined we have to make specific assumptions about the information that the agents have on each other's preferences. In this paper we investigate the consequences of two extreme assumptions. At one extreme we assume that every agent has no

---

[1] Actually Maskin did prove in [7] that only dictatorial voting schemes and voting schemes among two alternatives are such that for every profile the same outcome results from every Nash equilibrium of the associated game.

0022-0531/81/030398-15$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

information about any other agents' preference. This leads him to a *prudent* behaviour of the maximin type. At the other extreme we assume that every agent has full information about the preference profile; then, non-cooperation results in the *sophisticated* behaviour where the agents mutually anticipate their strategy by sucessively eliminating dominated strategies (see [2, 8, 9]).

To assume that everybody knows everybody elses utility does not avoid the implementation problem. The central legislator does not know the profile (at least when he must a priori choose a voting scheme that works for every profile); therefore in a world where no coercive device can prevent the individual agents from lying[2] decentralization must cope with the traditional incentive compatibility requirement, even if information about utility profile is complete.

The paper is organized as follows. In Section 1 we define prudent and sophisticated voting behaviour and the corresponding implementation of social choice functions (correspondences). In Section 2 we illustrate these concepts on a particular class of voting schemes, namely, voting by veto, in which we compare the prudent and sophisticated voting behaviours. The results stated in Section 2 are preliminary results for the next section. The main result is presented in Section 3. Namely, we exhibit a voting scheme, shortly described as voting by alternating veto, where both the prudent and the sophisticated voting of the agents yield the selection of the same alternative. The dictatorial voting schemes also share the property that the sophisticated voting and the prudent voting can not be distinguished. Thus, our result is worthwhile because voting by alternating veto in general is nearly anonymous (symmetrical among the players) and actually implements an exactly anonymous social choice function in some significant cases. These highly remarkable strategical properties of voting by veto makes from this familiar procedure a decision mechanism that can be defended on strong theoretical grounds (see also Ref. [10]). Finally Section 4 is devoted to some open problems and concluding remarks.

1. PRUDENT AND SOPHISTICATED VOTING

Throughout the paper we will consider a collective decision problem of the following form: the set A of alternatives is finite with cardinality p. Each of the n agents has a preference *ordering* on A (we assume that no agent is indifferent between any two alternatives; this assumption is of crucial technical importance). Let U be the set of strict ordering on A. For the sake of convenience we denote the current element u of U as a utility function—actually a one-to-one mapping from A into $\mathbb{R}$. All concepts and results will be purely ordinal.

---

[2] If the freedom-of-speech principle prevails then lying becomes even an individual right.

A profile u is an n-tuple $u = (u_1, \dots, u_n) \in U^n$ specifying the particular utility function of each agent.

DEFINITION 1. An n-person voting scheme among A is an $(n+1)$-tuple $\mathcal{V} = (X_1, \dots, X_n, \pi)$, where the strategy spaces, $X_1, \dots, X_n$ are finite and π is a mapping from $X_1 \times \cdots \times X_n$ into A.

DEFINITION 2. To every profile $u \in U^n$, the n-person voting scheme $\mathcal{V}$ among A associates the game in strategic form.

$$\mathcal{V}(u) = (X_1, \dots, X_n; u_1 \circ \pi, \dots, u_n \circ \pi).$$

In game $\mathcal{V}(u)$ we define now the prudent and sophisticated behaviour of the agents.

Notation. If $z = (z_1, \dots, z_K)$ is a vector of $\mathbb{R}^K$ we denote by $\theta(z)$ the vector obtained by reordering the coordinates of z in increasing order:

$$\theta(z) = (z_{\sigma(1)}, \dots, z_{\sigma(K)}), \qquad \begin{array}{l}\text{where } \sigma \text{ is a permutation of } \{1, \dots, K\} \\ \text{and } z_{\sigma(1)} \leqslant z_{\sigma(2)} \leqslant \cdots \leqslant z_{\sigma(K)}.\end{array}$$

DEFINITION 3. Let $\mathcal{V}$ be an n-person voting scheme among A and $u \in U^n$ be a particular profile. In the associated game $\mathcal{V}(u)$ we define the set $P_i(u_i)$ of ***prudent*** strategies of agent i as the set of those $x_i \in X_i$ which lexicographically maximize over X_i the vector $\theta\{u_i(x_i)\} = \theta\{u_i(x_i, x_{\hat{\imath}})_{x_{\hat{\imath}} \in X_{\hat{\imath}}}\}$.

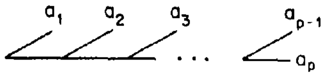
In particular a prudent strategy $x_i \in P_i(u_i)$ is a maximin strategy of player i

$$\min_{x_{\hat{\imath}} \in X_{\hat{\imath}}} u_i(x_i, x_{\hat{\imath}}) = \max_{y_i \in X_i} \min_{x_{\hat{\imath}} \in X_{\hat{\imath}}} u_i(y_i, x_{\hat{\imath}}).$$

The prudent behaviour of an agent is a decentralized one which is relevant when this agent has no information at all about the other agent's utility and consequently about their strategical choice. In many familiar voting schemes, like the plurality voting and the Borda count,[3] the prudent behaviour is nothing but the sincere voting (to announce his true peak alternative in plurality voting or his true preference ordering in the Borda procedure. The proof of these claims is left as an elementary exercise to the reader). In some other usual voting, like voting by binary choice, the notion of sincerity would not be enough to determine the prudent behaviour. Consider, for instance, the voting by successive amendments:

---

[3] Plurality voting and the Borda mehod permit ties. Therefore some tie-breaking device must be choosen before we can speak of voting schemes in the sense of Definition 1. Since this choice does not afect our statement we simply omit it.



By majority voting the agents decide first to elect a_1 or to reject it. If a_1 is rejected then the election of a_2 is proposed, and so on. At $(p-1)$th round of voting (if any) the agents select, again by majority voting a_p or a_{p-1}. Clearly the prudent strategy of agent i requires him to vote for a_k at the kth round unless he prefers a_k less than every following alternative $a_{k+1},\ldots, a_p$. This behaviour very much favors the first ranked alternatives! On the other hand, we must be aware that there ae some voting schemes where the agents have several prudent strategies that lead to the election of different alternatives. Consider, for instance, the following voting by veto procedure:

> Player 1 vetoes first one alternative, next player 2 vetoes one of the remaining alternatives, next player 1 vetoes again one of the remaining alternatives, next player 3 is a dictator to select one of the remaining alternatives.

In any of his prudent strategies, player 1 must decide to eliminate for sure his two least preferred alternatives. But he has two prudent strategies: one is to eliminate first his least preferred alternatives and next—after player 2's move—his least preferred among the $(p-2)$ remaining alternatives. The other prudent strategy is to eliminate first the alternative he ranks $(p-1)$ and next, after player 2's move, his least preferred among the $(p-2)$ remaining alternatives. One checks that the corresponding vectors $\theta(u_1(x_1))$ coincide, so that player 1 can not distinguish these strategies in terms of prudence. However, these two strategies may clearly yield the election of distinct alternatives.

DEFINITION 4. Let $\mathcal{V}$ be an n-person voting scheme among A. We will say that $\mathcal{V}$ p-implements (p stands for prudence) the following social choice correspondence SP

$$U^n \ni u \to SP(u) = \pi(P_1(u_1) \times \cdots \times P_n(u_n)) \subset A.$$

We turn now to the sophisticated behaviour of the agents.

Notation. Let $\mathcal{V} = (X_1,\ldots, X_n; \pi)$ be an n-person voting scheme among A and let $u = (u_1,\ldots, u_n) \in U^n$ be a fixed preference profile. The associated successive elimination of dominated strategies is the decreasing sequence $(X_1^t,\ldots, X_n^t)$ defined inductively as follows:

$$X_1^0 = X_1,\ldots, X_n^0 = X_n$$

for every $t = 0, 1, 2, \ldots$ and every $i \in \{1, \ldots, n\}$, X_i^{t+1} is the set of undominated strategies of agent i in the normal form game $(X_1^t, \ldots, X_n^t; u_1 \circ \pi, \ldots, u_n \circ \pi)$. That is to say

$$X_i^{t+1} = \{x_i \in X_i^t / \nexists y_i \in X_i^t\} \left\{ \begin{matrix} \forall x_{\hat{i}} \in X_{\hat{i}}^t u_i \circ \pi(x_i, x_{\hat{i}}) \leqslant u_i \circ \pi(y_i, x_{\hat{i}}) \\ \exists x_{\hat{i}} \in X_{\hat{i}}^t u_i \circ \pi(x_i, x_{\hat{i}}) < u_i \circ \pi(y_i, x_{\hat{i}}) \end{matrix} \right\}.$$

By definition, the sequence $(X_i^t)_{t=1,2,\ldots}$ is decreasing for every i, hence by the finiteness of X_i the sequence $(X_1^t, \ldots, X_n^t)$ is stationary for t large enough.

DEFINITION 5. We say that $\mathcal{V}$ is *dominant solvable* if for every profile $u = (u_1, \ldots, u_n) \in U^n$ the associated successive elimination of dominated strategies satisfies:

$$\text{for some } t,\ \pi(X_1^t \times \cdots \times X_n^t) \text{ contains a single alternative.} \tag{1}$$

In this case we will say that X_i^t is the set of sophisticated strategies of player i and denote it by $SO_i(u)$. Moreover we say that $\mathcal{V}$ d-implements (where d stands for dominance) the following social choice function SD:

$$U^n \ni u \to SD(u) = \pi(SO_1(u) \times \cdots \times SO_n(u)) \in A.$$

Let us briefly comment on the concept of dominance-solvable voting schemes (introduced first by Farqharson [2] and next studied in [8, 9]). This concept is intended to describe the voting behaviour of the agent being both completely informed of the utility profile and unable to cooperate in any way. Therefore, they mutually anticipate their behaviour by successively eliminating dominated strategies. The set $SO_1(u) \times \cdots \times SO_n(u)$ of sophisticated n-tuples of strategies (if it exists, that is, if (1) holds true) is a subset of the Nash equilibrium set and contains every equilibrium made up of dominating strategies (if any). Roughly speaking, the alternative selected by sophisticated voting of the players is a single-valued selection among the alternative selected by the various Nash equilibriums.

The class of dominance-solvable voting schemes is large. Essentially every voting scheme defined as a game in extensive form with perfect information is d-solvable, and the sophisticated voting is simply the perfect equilibrium of the game.[4] Thus voting by binary choice, voting by veto, and the king-maker procedure are d-solvable, whereas plurality voting or the Borda procedure are not. Given a particular d-solvable voting scheme, it p-implements a social choice correspondence SP and it d-implements a social choice function SD. To compare SD and SP amounts to comparing the sophisticated and prudent behaviour of the agents. This will throw some light on the kind of manipulation involved in a particular voting scheme. In Section 2 we describe *SD* and *SP* in the particular class of voting by veto procedure where these two social choice functions are different and in some sense symmetrical. Next in Section 3 we concentrate on voting schemes where the social choice function *SD* is a selection of the social choice correspondence *SP*; that is to say where the sophisticated voting can not be distinguished from a prudent behaviour.

---

[4] This fact was already suggested by Farqharson in [2]. It is proved in [8].

2. SOPHISTICATED AND PRUDENT VOTING BY VETO

DEFINITION 6. Let $p_1,\dots, p_n$ be non-negative integers such that $p_1 + \cdots + p_n = p - 1$ (remember that p is the cardinality of A). We denote by $\mathscr{T}[p_1,\dots, p_n]$ the n-person voting scheme among A defined as follows:

First player 1 vetoes any p_1 alternatives he wishes among A. Denote $A_1 \subset A$ the set of alternatives that he vetoes.

Next player 2 vetoes any p_2 alternatives among $A \backslash A_1$. Denote A_2 the set of alternatives he vetoes.

Then player 3 vetoes any p_3 alternatives among $A \backslash \{A_1 \cup A_2\}$.

And so on. If we denote by A_k the p_k alternatives vetoed by player k then the selected alternative is the single element of $A \backslash \{A_1 \cup \cdots \cup A_n\}$.

LEMMA 1. *In the voting scheme* $\mathscr{T}[p_1,\dots, p_n]$ *the prudent behavior of the players selects an unambiguous alternative: in other words, the social choice correspondence* $SP[p_1,\dots, p_n]$ *that this voting scheme p-implements actually is single valued.*

To prove Lemma 1 it suffices to describe the prudent strategies of the game $\mathscr{T}[p_1,\dots, p_n](u)$ for every profile $u \in U^n$.

Clearly player 1's unique prudent strategy is to veto his p_1 least preferred alternatives: it is actually his unique maximin strategy. Similarly, player 2's unique prudent strategy is to veto his p_2 least preferred alternatives among the $(p - p_1)$ remaining ones.

Thus the algorithm describing $SP[p_1,\dots, p_n]$ is as follows;

Let u be given and define inductively the subsets $A_1,\dots, A_n$ of A by

$$\begin{aligned} A_1 &= \text{the } p_1 \text{ least preferred alternatives of } u_1 \text{ among } A, \\ A_k &= \text{the } p_k \text{ least preferred alternatives of } u_k \text{ among} \\ &\qquad A\backslash\{A_1 \cup \cdots \cup A_{k-1}\}. \end{aligned} \tag{2}$$

Then we have

$$SP[p_1,\dots,p_n](u)=a \Leftrightarrow \{a\}=A\backslash\{A_1\cup\cdots\cup A_n\}.$$

LEMMA 2. *In voting scheme $\mathcal{V}'[p_1,\dots,p_n]$ the sophisticated behavior of the players selects the same alternative as the* prudent *behavior of the players does in the voting scheme $\mathcal{V}'[p_n,p_{n-1},\dots,p_1]$. That is to say, we have the formula*:

$$SD[p_1,p_2,\dots,p_n]=SP[p_n,p_{n-1},\dots,p_1]. \tag{3}$$

The above Lemma claims hat given the profile $u=(u_1,\dots,u_n)$ the sophisticated voting of the players is defined by the following algorithm:

$$\begin{aligned}&\text{Player } n \text{ vetoes the set } B_n \text{ of the } p_n \text{ least preferred}\\ &\quad\text{alternatives of } u_n \text{ among } A.\\ &\text{For } k=n-1,n-2,\dots,1, \text{ player } k \text{ vetoes the set } B_k \text{ of}\\ &\quad\text{the } p_k \text{ least preferred alternatives of } u_k \text{ among}\\ &\quad A\backslash\{B_n\cup B_{n-1}\cup\cdots\cup B_{k+1}\}.\end{aligned} \tag{4}$$

Then we have

$$SD[p_1,\dots,p_n](u)=a \Leftrightarrow \{a\}=A\backslash\{B_n\cup B_{n-1}\cup\cdots\cup B_1\}.$$

The sophisticated voting of player 1 when the profile is u is to veto the alternatives of B_1: in order to compute B_1 he must then compute the whole sequence $B_n, B_{n-1},\dots,B_2$, thus using his complete knowledge of the other players' preferences. This illustrates the high complexity of sophisticated voting.

Proof of Lemma 2. We prove by induction on n, the number of players, that the sophisticated voting in game $\mathcal{V}'[p_1,\dots,p_n](u)$ is described by algorithm (4). This claim is trivial for $n=1$. We assume it holds true up to $(n-1)$. We compute now the sophisticated strategy of player 1. If he vetoes some set C_1 of p_1 alternatives, then he can (by the inductive assumption) assume that the $(n-1)$ other players will successively veto the subset $C_2,\dots,C_n$ of A defined inductively as

$$\begin{aligned}&C_n \text{ is the set of the } p_n \text{ least preferred alternatives of } u_n\\ &\quad\text{among } A\backslash C_1.\\ &\text{For } k=(n-1),(n-2),\dots,2,\ C_k \text{ is the set of the } p_k\\ &\quad\text{least preferred alternatives of } u_k \text{ among}\\ &\quad A_1\backslash\{C_1\cup C_n\cup C_{n-1}\cup\cdots\cup C_{k+1}\}.\end{aligned} \tag{5}$$

The selected alternative is finally c defined by

$$\{c\}=A\backslash\{C_1\cup C_n\cup C_{n-1}\cup\cdots\cup C_2\}.$$

We must now prove that c is not preferred by u_1 to a defined by (4). For that purpose we compare the sequences B_n, B_{n-1},... and C_n, C_{n-1},..., by which a and c have been successively defined. We remark that if B and C are any two subsets of A such that $B \subset C$ and if we set

$$\tilde{B} = \text{the } p_k \text{ least preferred alternatives of } u_k \text{ among } A \backslash B,$$

$$\tilde{C} = \text{the } p_k \text{ least preferred alternatives of } u_k \text{ among } A \backslash C,$$

then we have $\tilde{B} \subset C \cup \tilde{C}$.
Applying this remark to $B = \phi \subset C_1 = C$ and $k = n$ we obtain

$$B_n \subset C_1 \cup C_n.$$

Applying the remark again to the above inclusion and $k = (n-1)$, we obtain

$$B_{n-1} \subset C_1 \cup C_n \cup C_{n-1} \Rightarrow B_n \cup B_{n-1} \subset C_1 \cup C_n \cup C_{n-1}.$$

Clearly we can apply the remark inductively for decreasing k, so that for every k we obtain

$$B_n \cup \cdots \cup B_k \subset C_1 \cup C_n \cup \cdots \cup C_k.$$

and finally

$$B_n \cup \cdots \cup B_2 \subset C_1 \cup C_n \cup \cdots \cup C_2,$$

that is to say,

$$c \in A \backslash \{B_n \cup \cdots \cup B_2\} = B_1 \cup \{a\}.$$

By definition of B_1 we have

$$\forall a' \in B_1 \cup \{a\}, \qquad u_1(a) \succcurlyeq u_1(a').$$

Therefore $u_1(a) \succcurlyeq u_1(c)$, which was to be proved. By Lemmas 1 and 2 the sophisticated and prudent behaviors of the players in $\{p_1, \ldots, p_n\}$ are both deterministic and are deduced from one another by simply reversing the ordering of the players.

3. Voting by Alternating Veto

Voting by veto procedures introduce a strong disymmetry among agents: even if the *veto powers* $p_1, \ldots, p_n$ are all equal, or nearly equal, the ordering of the agents has a strong influence on the outcome of both the sophisticated and the prudent voting (this obvious intuition is made clear by the two algorithms above, (2) and (4)).

To reduce and sometimes avoid this dissymmetry we will replicate the alternatives and at the same time alternate the strategical vetoes of the players.

Definition 7. Let r and, for every alternative a, k_a be non-negative integers such that

$$a - r = \left(\sum_{a \in A} k_a\right) - 1. \tag{6}$$

To every r, k_a, verifying (6), we associate a family of procedures defined as follows:

Each agent is endowed with r tokens. There are r rounds of vetoing. At each round the agents are successively asked (the ordering of the agents can change at each round, but the agents cannot influence it) to throw one token over one of the alternatives, a, on which at most $(k_a - 1)$ tokens have already been thrown. The elected alternative is the alternative a which after r rounds received only $(k_a - 1)$ tokens (every other alternative b having in view of (6) received exactly k_b tokens). We will call these procedures voting by alternating veto. This terminology is made clear by the following interpretation of the above procedure: each alternative is replicated a certain number of times, alternative a being replicated k_a times. At each round the players must successively veto one replica of one of the remaining alternatives.

The basic property of voting by alternating veto is that if we order carefully the successive vetoes by the agents, then the sophisticated behaviour of the agents cannot be distinguished from a prudent one.

Definition 8. Let $\mathcal{V} = (X_1, \dots, X_n, \pi)$ be a d-solvable n person voting scheme among A. We say that $\mathcal{V}$ is *exactly* solvable if for every profile $u \in U^n$, the sophisticated voting of the players also is a prudent voting. That is to say,

$$\forall u \in U^n SD(u) \in SP(u). \tag{7}$$

If (7) holds true we say that V *exactly* implements the social choice function SD.

This terminology is derived from Peleg's concept of an exactly consistent voting scheme (see [11]). In an exactly consistent voting scheme, there exists for every profile an equilibrium n-tuple which cannot be distinguished from the "tell-the-truth" strategy. However, there is no obvious comparison of the two concepts.

A strategy-proof voting scheme is in particular exactly solvable since both the sophisticated and the prudent behaviour of a player amount to selecting a dominant strategy. In general exact-solvability is a weakening of strategy-proofness which amounts to saying that the completely informed, non-cooperative agents eventually select the same alternative as completely uninformed risk-averse agents. The next theorem displays a family of exactly solvable voting schemes.

THEOREM 1. *Suppose r is an even integer and k_a, $a \in A$ are non-negative integers verifying* (6). *Then the following associated voting by veto procedure is exactly solvable. During the first $r/2$ rounds, the ordering of the players is fixed, say $1,\dots, n$. During the last $r/2$ rounds this ordering is simply reversed, say n, $(n-1),\dots, 1$.*

Proof. We first introduce a useful notation: for every finite sequence $i_1,\dots, i_{q-1}$ with values in the set $\{1,\dots, n\}$ of agents, we denote by $W[i_1,\dots, i_{q-1}]$ the n-person voting scheme among q alternatives (where q will be in fact much greater than p) in which the players successively veto one of the remaining alternatives. More precisely the procedure works as follows: player i_1 vetoes first one alternative, say a_1; the next player i_2 vetoes one alternative, say a_2 among $A - \{a_1\}$; and so on; at stage t, player i_t vetoes one alternative, say a_t, among $A - \{a_1,\dots, a_{t-1}\}$; because a contains q elements, there is exactly one alternative a_q in $A - \{a_1,\dots, a_{q-1}\}$: this is the elected alternative.

For instance, the voting scheme $\mathcal{V}[p_1,\dots, p_n]$ is now equivalently written as

$$W[\underbrace{1,\dots,1}_{p_1 \text{ times}}, \underbrace{2,\dots,2}_{p_2 \text{ times}},\dots, \underbrace{n,\dots,n}_{p_n \text{ times}}].$$

In the voting scheme described in the statement of Theorem 1 there are $q = \sum_{a \in A} k_a$ alternatives: alternative a is replicated k_a-times. The procedure is then

$$W[\underbrace{1\,2,\dots,n}_{1}, \underbrace{1,\dots,n}_{2},\dots, \underbrace{1,\dots,n}_{r/2}, \underbrace{n(n-1),\dots,1}_{1}, \underbrace{n,\dots,1}_{2},\dots, \underbrace{n,\dots,1}_{r/2}]. \tag{8}$$

In order to prove the theorem we must describe the sophisticated voting of the players in (8) and check that it coincides with a prudent behaviour. In view of algorithm (4) describing the sophisticated voting in the voting scheme $\mathcal{V}[p_1,\dots, p_n]$ we obtain easily the sophisticated voting in $W[i_1,\dots, i_{q-1}]$. Denote by B the set of q (different) alternatives, and $u = (u_1,\dots, u_n)$ a particular profile. Then $SD(u)$ is described by the following algorithm.

Let b_{q-1} be the least preferred alternative of $u_{i_{q-1}}$ among B.
Let b_{q-2} be the least preferred alternative of $u_{i_{q-2}}$ among $B\backslash\{b_{q-1}\}$.
· · ·
Let b_k be the least preferred alternative of u_{i_k} among $B\backslash\{b_{q-1},\ldots, b_{k+1}\}$.
· · ·
Then $b = SD(u) \Leftrightarrow \{b\} = B\backslash\{b_{q-1},\ldots, b_1\}$. (9)

Remark that in order that each alternative b_k is well defined we have assumed that each utility function is one-to-one over B. Actually we can weaken this assumption and still give sense to algorithm (6). For instance, suppose that if one agent is indifferent between two alternatives of B, then every other agent also is indifferent:

$$\forall i, j \in \{1,\ldots, n\}\ \forall a, b \in B(u_i(a) = u_i(b)) \Leftrightarrow (u_j(a) = u_j(b)). \tag{10}$$

Then two alternatives a, b such that every agent is indifferent between them can be identified so that algorithm (9) yields undiscernable alternatives. In the case of voting scheme (8), the set B has the form

$$B = \{\underbrace{a,\ldots, a}_{k_a \text{times}}, \underbrace{b,\ldots, b}_{k_b \text{times}},\ldots\}.$$

By assumption the utility functions u_i are one-to-one on A so that on B property (10) holds. Therefore a straightforward application of algorithm (9) yields the sophisticated voting in scheme (8).

On the other hand, we remark that in voting scheme $W(i_1,\ldots, i_{q-1})$ it is a prudent behaviour for any player to veto his least preferred alternative among the remaining ones whenever he has to veto some alternative. Thus the following algorithm defines a selection of the social choice correspondence p-implemented by $W(i_1,\ldots, i_{q-1})$:

Let b_1 be the least preferred alternative of u_{i_1} among B.
Let b_2 be the least preferred alternative of u_{i_2} among $B\backslash\{b_1\}$.
· · · · (11)
Let b_k be the least preferred alternative u_{i_k} among $B\backslash\{b_1,\ldots, b_{k-1}\}$.
Then if $\{b\} = B\backslash\{b_1,\ldots, b_{q-1}\}$, we have $b \in SP(u)$.

Comparing algorithm (9) and (11) it is clear that $W(i_1,\ldots, i_{q-1})$ is exactly solvable if the two sequences $\{i_1, i_2,\ldots, i_{q-1}\}$ and $\{i_{q-1}, i_{q-2},\ldots, i_2, i_1\}$ coincide. It is clearly the case for the voting scheme (8); this completes the proof of Theorem 1.

The property of exact-solvability is not by itself sufficient to make a voting scheme ethically attractive. An exactly solvable voting scheme is nothing but a procedure wherein the collective implications of non-cooperation are easily predictable, just as they are in a strategy-proof dictatorial voting scheme. To make the voting scheme more attractive, we would like that the social choice function that it exactly implements shares some additional properties. Among these, the three most usually desired are efficiency, anonymity and neutrality. We say that the social choice function S is

efficient if $S(u)$ is a Pareto optimum alternative of the profile u;

anonymous if $S(u_1,\dots,u_n)$ is a symmetrical function of $(u_1,\dots,u_n)$ (S does not discriminate among players);

neutral if S does not discriminate among alternatives, which is formalized as follows: for every one-to-one mapping σ from A into itself, and every profile $(u_1,\dots,u_n)$:

$$\sigma(S(u_1 \circ \sigma,\dots, u_n \circ \sigma)) = S(u_1,\dots,u_n).$$

LEMMA 3. *For every r, k_a ($a \in A$) verifying* (6) *every associated voting by alternating veto d-implements an efficient social choice function.*

Proof. It suffices to show that in any voting scheme of the form $W[i_1,\dots,i_{q-1}]$ the alternative b selected by algorithm (9) is efficient. By (9) every alternative other than b is one of the b_k's and therefore is less preferred than b by at least one utility function u_{i_k}. Therefore it cannot dominate b.

LEMMA 4. *The social choice function exactly implemented by the voting by alternating veto* (8) *is neutral if and only if $k_a = k$ does not depend on a. This is possible only if p is odd and relatively prime with respect to n.*

Proof. Suppose r, k_a ($a \in A$) are given such that k_a does depend on a. Then we fix $a, b \in A$ such that $k_a < k_b$. In this case we prove that any corresponding voting by alternating veto implements a non-neutral social choice function. Suppose that the profile $(u_1,\dots,u_n)$ is such that:

$$\forall i = 1,\dots, n, \qquad \forall c \in A, \qquad c \neq a, \qquad b \Rightarrow u_i(c) < \inf\{u_i(a), u_i(b)\}.$$

Then we set $k = k_a + k_b - 1$ and we develop our particular voting by alternating veto which can be written as

$$W[i_1,\dots, i_k, i_{k+1},\dots, i_{n\cdot r}].$$

In view of the profile that we consider, the sophisticated voting of the players, described by algorithm (9), guarantees that every player after stage $k+1$ will veto one replica of one alternative of $A\setminus\{a, b\}$. Since $n \cdot r - k = \sum_{c \in A\setminus\{a,b\}} k_c$ we deduce that at every stage before stage k (including stage k) the players will veto one replica either of a or of b. We can then select a profile such that from i_1 to i_k, the players who prefer a more than b are as many as the players who prefer b more than a, or these two numbers differ at most by one:

$$|\#\{j=1,\ldots,k/U_{i_j}(a)>U_{i_j}(b)\}-\#\{j=1,\ldots,k/U_{i_j}(b)>U_{i_j}(a)\}|\leqslant 1. \quad (12)$$

Then one checks easily that alternative a is selected by the sophisticated voting of the players. If we consider the permutation σ: $\sigma(a)=b$; $\sigma(b)=a$; $\sigma(c)=c$ otherwise, then property (12) still holds true for the profile $(u_1\circ\sigma,\ldots,u_n\circ\sigma)$ so that alternative a is again selected by sophisticated players. This contradicts the neutrality of the social choice function implemented by this voting by alternating veto procedure.

Conversely, if k_a does not depend on a, it is obvious that the corresponding voting by alternating veto are neutral procedures. Coming back to equation $r\cdot n=\sum_{a\in A}k_a-1$ this yields

$$r\cdot n=k\cdot p-1. \quad (13)$$

An elementary result of arithmetic (known as the Bezout identity) yields the following statement: if n and p are given, then there exist non-negative integers r, k with r even, verifying (13) if and only if p is odd and relatively prime with respect to n.

We come now to the anonymity property; in our context it is the most desirable since the dictatorial social choice functions are bad only with respect to anonymity arguments. At this point the use of a rather complex procedure like the one we considered becomes fully justified. Notice first that voting by alternating veto procedures are essentially anonymous because every agent is endowed with the same number of tokens, that is to say, the same "veto power."[5] Moreover, when r, the number of rounds, becomes large, intuition suggests that the ordering of the players becomes less relevant so that the implemented social choice function becomes nearly anonymous. This intuition actually is a rigorous statement for some value of n:

LEMMA 5. *Let n be a prime integer greater than p. Let r and k be non-negative integers such that*

$$r\cdot n=k\cdot p-1. \quad (14)$$

Then for r, k large enough verifying (14), *the voting by alternating veto procedures where $k_a=k$ for every a d-implement an anonymous social choice function. The proof of this Lemma, which is by no means trivial, can be found in* [9].

---

[5] Notice that when $r<k$, i.e., when $p<n$, no agent can alone veto even one alternative.

Gathering the results of the above three Lemmas, we obtain the following theorem which summarizes the desirable properties of voting by alternating veto.

THEOREM 2. *Let n be a prime integer, let p be odd and strictly smaller than n. Then there exists a pair r, k of non-negative integers verifying*

$$r \cdot n = k \cdot p - 1$$

and an associated voting by alternating veto procedure, which exactly implements an efficient, anonymous and neutral social choice function.

The arithmetic condition (n should be prime and greater than p) is very natural since there does not exist an efficient anonymous and neutral social choice function unless every prime factor of n is greater than p. Let us mention finally that the social choice function exactly implemented by alternating veto is also a monotonic social choice function (this is obvious in view of algorithm (9)), a property shared by most of the social choice functions implemented by dominance solvable voting schemes.

4. CONCLUSION AND OPEN PROBLEMS

The very concept of exact implementation hopefully should be strengthened to make its strategical interpretation more convincing. Instead of requiring that $SD(u)$, the alternative selected by the sophisticated voting of the agents, should be a member of $SP(u)$, the possibly numerous alternatives that can be selected by their prudent behaviour, it would be undoubtedly more convincing to demand $SD(u) = SP(u)$, which entails the additional requirement that the prudent behaviour actually is non-ambiguous. This is not satisfied by voting by alternating veto since the prudent behaviour might be ambiguous, even for large values of r.

Suppose for instance $p = 3$ and $n = 5$, and choose a pair of (large) integers r, k such that

$$5r = 3k - 1.$$

Suppose agent 1's preferences are

$$u_1(a) > u_1(b) > u_1(c).$$

Because $r < k$ player 1 cannot a priori prevent c from being elected. Therefore his prudent behaviour during the first rounds of vetoing is to systematically veto one replica of alternative c. Suppose now that after t rounds of vetoing, $(k - r + t + 1)$ tokens have been thrown on c by player 1 *himself as well as by the other players. Because player 1 is left with $r-t$ tokens he has now the power to prevent the election of alternative c and still being left with one token; namely:*

$$(r-t)+(k-r+t+1)=k+1.$$

Accordingly he has several options within his very prudent behaviour: he can decide to throw first one token over b *and then his* $(r-t-1)$ *remaining* tokens over c if nobody else contributes to the elimination of c; alternatively he can decide to prevent as soon as possible the election of c. These two options are not equivalent and this explains why $SP(u)$ is not in general single valued. The author failed to determine whether or not Theorem 1 (as well as Theorem 2) still holds true for some reasonable family of voting procedures, if we use the stronger version of exact implementation mentioned above. We feel however that Theorem 2 allows us to support voting by alternating veto as a highly valuable voting scheme both with respect to its strategical and equity properties.

REFERENCES

1. B. Dutta and P. K. Pattanaik, On nicely consistent voting systems, *Econometrica* **46** (1978), 163–170.
2. R. Farqharson, "Theory of Voting," Yale Univ. Press, New Haven, Conn., 1969.
3. J. Green and J. J. Laffont, Characterization of satisfactory mechanism for the revelation of preferencers for public goods, *Econometrica* **45**, 2 (1977).
4. T. Groves, "Efficient Collective Choice with Compensation," Technical report No. 258 IMSSS, Stanford University, 1978.
5. L. Hurwicz and D. Schmeidler, Construction of outcome functions guaranteeing existence and Pareto optimality of Nash equilibria, *Econometrica* **46**, No. 6 (1978), 1447–1474.
6. E. Maskin, Nash equilibrium and welfare optimality, *in* "Mathematics for Operations Research," in press.
7. E. Maskin, "Implementation and Strong Nash Equilibrium," MIT Working paper No. 216, 1978.
8. H. Moulin, Dominance solvable voting schemes, *Econometrica* **47**, No. 6 (1979), 1337, 1352.
9. H. Moulin, Implementing efficient, anonymous and neutral social choice functions, *J. Math. Econ.* 7 (1980), 249–269.
10. D. Mueller, Voting by veto, *J. Public Econ.* **10**, No. 1 (1978), 57–76.
11. B. Peleg, Consistent voting systems, *Econometrica* **46**, No. 1 (1977), 153.

Two-Sector Model with Quantity Rationing*

CLAUDE FOURGEAUD

Université Paris I, Centre de Mathématiques
Économiques, Paris 75005, France, and
CEPREMAP, 142 rue du Chevaleret, Paris, France

BERNARD LENCLUD

Ministère de l'Économie, Direction de la Prévision,
151 rue Saint-Honoré, Paris 75001, France

AND

PHILIPPE MICHEL

Université Paris I, Centre de Mathématiques Économiques,
Paris 75005, France

Received November 14, 1979; revised July 11, 1980

INTRODUCTION

The analysis of short-term economic equilibrium based on the adjustment of quantities, prices being fixed, has developed considerably in recent years. Parallel to formalizations in terms of general equilibrium and of reformulation of the microeconomic foundations of macroeconomy ([3, 6, 12, 14]; see the survey in [7]), an aggregate macroeconomic model has been the object of systematic studies. Following Clower's article on the dual decision, Barro and Grossman [1] bring out the two situations involving excess supply and excess demand on the two markets (labour, goods) according to the level of prices and the level of wages. Benassy [4] shows the possibility of a third area, that of classical unemployment. Various studies have developed the basic model ([2, 9, 11]): Malinvaud, in particular, emphasizes in the appendix to [11] the existence of a fourth area in case of stockage, an area dealt with in [13]. These analyses have made it possible to lay foundations for macroeconomic policies with greater precision, despite the schematic nature of the model, which comprises only one sector of production. Nevertheless, the corresponding economic policies are overly rational and

---

* This work was supported in part by Direction de la Prévision, Ministère de l'Économie, France.

0022-0531/81/030413-24$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved

comforting. A wide variety of situations may exist in different sectors of the economy. This has prompted us to try and ascertain whether less clear-cut consequences appear in a model which distinguishes at least two sectors, those of consumer goods and investment goods.

Our study presents an example of a two-sector model which takes into account both consumer and investment demand with differentiated behaviours in each sector. In addition to labour and money, the model comprises two final goods and one financial asset: bonds issued by the sector of production to finance their investments, and subcribed to by households. Public expenditures, financed by the creation of money, are both consumption and investment expenditures.

The logic of the model is essentially the logic of fixed-price models: the two prices of the goods and the wage rate are fixed. Conversely, the interest rate is the result of equilibrium between supply and demand on the securities market. Possible rationing in both sectors, on the investment goods and labour markets, is defined by a proportional scheme.

We shall prove the existence of an equilibrium for the system, analyzed according to the types of equilibrium prevailing in each of the two sectors: we thus obtain a typology by combining the various situations of partial equilibrium in each sector.

The variations in the parameters determining equilibrium are analyzed. The consequence of certain economic policies are not as clear-cut as they appear to be in a single-sector model. For instance, in the Keynesian situation, there is a conflict between policies emphasizing an increase in output and those emphasizing employment. The impact of a variation in nominal wages on activity depends both on the sectors themselves and their type of equilibrium. This also applies to the effect of the parameters on the rate of interest, such that the final effect of certain measures appears to be highly uncertain.

The study is in two parts. Part I provides a description of the different elements of the model and the study of equilibrium. Part II analyzes the effects of policies and parameters on the general equilibrium in three situations: these being the two situations in which Keynesian and classical unemployment prevail, plus the third situation, where both sectors are in Keynesian unemployment.

I. PRESENTATION OF THE MODEL AND STUDY OF THE EQUILIBRIUM

I.1. *The economy under consideration* comprises

A *capital good*, price p_1;
consumer good, price p_2;

labour, nominal wage rate: w;

money, price $p_0 = 1$;

bonds issued by the firms to finance their investments, and subscribed to by households (price q).

The four prices p_1, p_2, w and p_0 are assumed to be fixed; the equilibrium is a short-term one achieved by quantity rationing. On the other hand, it is assumed (as it was assumed by Hool [10]) that the price q of the securities can vary on a market of the open-market type to balance supply and demand.

The four agents of the economy under consideration are

Households;

government;

the *capital goods sector of production* (*sector* 1);

the *consumer goods sector of production* (*sector* 2).

Notation. Equilibrium values are expressed in barred symbols.

I.2. The households's characteristics are

The *initial stock of money*: $M_{c,0}$;

the *initial stock of securities*: B_0;

their *current incomes*: $D_1(\bar{L}_1) + D_2(\bar{L}_2)$, which are the result of employment levels at equilibrium $\bar{L}_i$, in sector i ($i = 1, 2$).

The following assumptions are made:

1. the labour supply L^0 is constant.

2. Aggregate propensity to consume is constant: with wealth $R(\bar{L}_1, \bar{L}_2)$, aggregate effective demand for consumption good C^d verifies

$$p_2 C^d = \alpha R(\bar{L}_1, \bar{L}_2), \qquad 0 < \alpha < 1. \tag{1}$$

3. There exists a linear relation between aggregate demands for bonds and for money:

$$\beta q(B + Q) - \gamma(M_c + P) = 0, \qquad \beta > 0, \gamma > 0, \tag{2}$$

where constants P and Q are related to expected income for subsequent periods. We assume $\alpha + \beta + \gamma = 1$, $\alpha > \beta$ and $\alpha > \gamma$.[1]

The *budgetary constraint* of households is

$$p_2 C + M_c + \bar{q}B = R(\bar{L}_1, \bar{L}_2). \tag{3}$$

---

[1] These properties are obtained if one assume that aggregate behaviour of households may be defined by the maximization of a Cobb–Douglas utility function (which does not depend on labour): $U = \alpha \log C + \beta \log(P + M_c) + \gamma \log(Q + B)$.

We shall see that for any employment levels L_1 and L_2 in the two sectors, we can derive from the equality between the supply of and demand for securites, an equilibrium price $q(L_1, L_2)$. Thus, the wealth of households is associated with those levels of employment:

$$R(L_1, L_2) = M_{c,0} + q(L_1, L_2)\,B_0 + D_1(L_1) + D_2(L_2) \tag{4}$$

and we have, for the short-term equilibrium: $\bar{q} = q(\bar{L}_1, \bar{L}_2)$.

I.3. *The government*'s activities are confined to:

a given consumption G;
a given investment I_G;
the creation of money corresponding to payment for its consumption and for its capital goods investment:

$$\Delta M = p_1 I_G + p_2 G. \tag{5}$$

I.4. *The consumer goods sector.* The production function of sector 2, $F_2(L_2)$ depends upon its level of employment L_2, and upon a production capacity K_2, which is not explicitly considered for the short run equilibrium. This sector maximizes its profit

$\Pi_2(L_2) = p_2 F(L_2) - wL_2$ subject to the following constraints:

$$\begin{aligned} \text{employment}: L_2 &\leqslant L^0 - \bar{L}_1, \\ \text{production outlets}: F_2(L_2) &\leqslant C^d + G, \end{aligned} \tag{6}$$

C^d being the effective demand of households.

Sector 2 has a given initial stock of money $M_{2,0}$ which is assumed to be net of interest on the securities $B_{2,0}$. It decides to keep cash in hand: M_2 and it distributes income $D_2(L_2)$. The balance of its assets is

$$r_2(L_2) = p_2 F_2(L_2) - D_2(L_2) + M_{2,0} - M_2. \tag{7}$$

These resources (which are generally positive) are completed by the issue or the redeeming of securities: $B_2 - B_{2,0}$, to finance its investment I_2: $p_1 I_2 = r_2(L_2) + q(B_2 - B_{2,0})$.

The investment which is actually realized is the minimum of (1) the demand for capital good I_2^d and (2) the limit to physical acquisitions possibilities:

$$I_2 = \min\{I_2^d, F_1(\bar{L}_1) - I_G - I_1\}; \tag{8}$$

and the demand I_2^d is a given function A_2 of the price q of securities and various parameters which are constant in the short run: K_2, $p_1, p_2, w, \ldots,$

$$I_2^d = A_2(q). \tag{9}$$

I.5. *The capital goods sector* (sector 1). Sector 1 has a production function $F_1(L_1)$.

It maximizes its profit: $\Pi_1(L_1) = p_1 F_1(L_1) - wL_1$ subject to the following constraints:

$$\begin{aligned} &\text{employment: } L_1 \leqslant L^0 - \bar{L}_2, \\ &\text{production outlets, } F_1(L_1) \leqslant I_1^d + I_2^d + I_G. \end{aligned} \tag{10}$$

The investment demanded and realized in sector 1 are defined in the same way as in sector 2:

$$\begin{aligned} I_1^d &= A_1(q), \\ I_1 &= \min\{I_1^d, F_1(L_1) - I_G - I_2\}. \end{aligned} \tag{11}$$

The initial stock of money $M_{1,0}$ is given; the final stock decision is M_1, and the immediately distributed income is $D_1(L_1)$. The investment realized is financed by the balance $r_1(L_1)$ and by the issuing (or withdrawal) of securities: $p_1 I_1 = r_1(L_1) + q(B_1 - B_{1,0})$. $r_1(L_1)$ is the balance: $r_1(L_1) = p_1 F_1(L_1) - D_1(L_1) + M_{1,0} - M_1$.

I.6. *Quantity rationing.* The *securities markets* is balanced, and subject to no constraints other than: $B_1 \geqslant 0$ and $B_2 \geqslant 0$.

The *money market* is also cleared. If we add the physical equilibrium equations, expressed in current prices to the budgetary equilibrium equations, we obtain:

$$p_2 G + p_1 I_G = \Delta\bar{M}_c + \Delta\bar{M}_1 + \Delta\bar{M}_2 = \Delta M, \tag{12}$$

which is precisely the definition of government-created money.

On the *consumer goods market*, it is assumed that the government is the first to be served, so that: $\bar{C} = \max\{0, F_2(\bar{L}_2) - G\}$. On the *capital good market*, the government is the first to be satisfied. Then the sectors are subjected to proportional rationing. On the *labour market*, the two sectors are subjected to proportional rationing.

I.7. *Equilibrium of the securities market.* Let us consider the levels of activity of the two sectors as defined by levels of employment L_1 and L_2, respectively. At these levels, by virtue of the budgetary constraint of the government, the *stock of money held by the households is*

$$M_c = p_2 G + p_1 I_G + M_{c,0} + M_{1,0} + M_{2,0} - M_1 - M_2. \tag{13}$$

The supply of securities in both sectors is defined by their budgetary equilibrium; by adding and taking into account the physical equilibrium of sector 1, we obtain

$$r_1(L_1) + r_2(L_2) + q(B - B_0) = p_1(I_1 + I_2) = p_1(F_1(L_1) - I_G), \quad (14)$$

and substituting r_1 and r_2, the supply of securities is such that

$$\begin{aligned} q(B - B_0) = D_1(L_1) + M_1 - M_{1,0} - (p_2F_2(L_2) - D_2(L_2) \\ - M_2 + M_{2,0}) - p_1 I_G. \end{aligned} \quad (15)$$

With relations (2) and (15) defining the securities demand, we obtain the equilibrium price of securities:

$$\begin{aligned} q(L_1, L_1) \cdot (B_0 + Q) = {} & (\gamma/\beta)(P + p_2 G + M_{c,0}) + (\gamma/\beta + 1)\, p_1 I_G \\ & - D_1(L_1) - (\gamma/\beta + 1)(M_1 + M_2 - M_{1,0} - M_{2,0}) \\ & + p_2F_2(L_2) - D_2(L_2). \end{aligned} \quad (16)$$

This relation may also be expressed as a function of M_c (relation (13)):

$$\begin{aligned} q(L_1, L_2) \cdot (B_0 + Q) = (\gamma/\beta)(P + M_c) - D_1(L_1) - D_2(L_2) \\ + M_c - M_{c,0} + p_2(F_2(L_2) - G). \end{aligned} \quad (17)$$

I.8. *General equilibrium.* Substituting the equilibrium value of q (relation (17)) into the household's wealth (relation (4)), the effective demand for consumption good (relation (1)) becomes

$$p_2 C^d = \alpha [p_2(F_2(L_2) - G) + (\gamma/\beta + 1)(P + M_c)]. \quad (18)$$

The following levels are defined for sector 2:

level $L_2^1(L_1)$ of full employment of labour: $L_2^1(L_1) = L^0 - L_1$;

level of employment L_2^2 corresponding to the maximum profit: $F_2'(L_2^2) = w/p_2$;

level of employment L_2^3 corresponding to the satisfaction of demand: $F_2(L_2^3) = G + C^d(L_2^3)$.

Thus, with the usual concavity assumptions on F_2, for a given L_1, *the equilibrium of the sector of the consumer good* corresponds to the level of employment:

$$\mathcal{L}_2(L_1) = \min\{L_2^1(L_1), L_2^2, L_2^3\} \quad (19)$$

and its demand for labour is

$$L_2^d = \min\{L_2^2, L_2^3\}. \quad (20)$$

Remark. Expression (19) defines three major types of equilibrium for sector 2 [11]: *repressed inflation* for $\hat{L}_2 = L_2^1$, *classical unemployment* for $\hat{L}_2 = L_2^2$, *Keynesian unemployment* for $\hat{L}_2 = L_2^3$.

For an employment level L_1 in sector 1, the investment demands are

$$\hat{I}_1^d = A_1(q(L_1, \hat{L}_2(L_1)) \qquad \text{and} \qquad \hat{I}_2^d = A_2(q(L_1, \hat{L}_2(L_1)). \tag{21}$$

Then the total demand for investment is

$$I^d(L_1) = \hat{I}_1^d(L_1) + \hat{I}_2^d(L_1) + I_G. \tag{22}$$

Assuming that F_1 satisfies the usual concavity assumptions, and that $F_1(L_1) - I^d(L_1)$ is an increasing function of L_1, the demand for labour L_1^d is defined by

$$L_1^d = \min\{L_1^2, L_1^3\}, \qquad \text{with} \quad F_1'(L_1^2) = w/p_1$$

and

$$F_1(L_1^3) = I^d(L_1^3). \tag{23}$$

The *equilibrium of sector* 1 corresponding to a level L_2 of employment in sector 2, is defined by: $\bar{L}_1 = \min\{L_1^d, L_0 - \bar{L}_2\}$.

A *fixed-price equilibrium* corresponds to equilibria for households, sector 1 and sector 2 such that: (1) on the markets for labor, consumption good and capital good, transactions realize the minimum of supply and demand; (2) the bonds market is cleared; (3) rationing verifies conditions of Section 1.6.

Let

$$\begin{aligned} &\text{if} \quad L_1^d + L_2^d \leqslant L^0, \qquad \bar{L}_1 = L_1^d \quad \text{and} \quad \bar{L}_2 = L_2^d, \\ &\text{if} \quad L_1^d + L_2^d > L^0, \qquad \bar{L}_1 = aL_1^d \quad \text{and} \quad \bar{L}_2 = aL_2^d, \\ &\qquad\qquad\qquad\qquad\qquad\qquad a = L^0/(L_1^d + L_2^d). \end{aligned} \tag{24}$$

It is assumed that: $G \leqslant F_2(\bar{L}_2)$ and $I_G \leqslant F_1(\bar{L}_1)$, i.e., the government demands are satisfied.

THEOREM. *For $\bar{L}_1$ and $\bar{L}_2$ defined by* (24), *the following values $\bar{C}$, $\bar{q}$, $\bar{I}_1$, $\bar{I}_2$, $\bar{B}_1$, $\bar{B}_2$*:

$$\begin{aligned} &\bar{C} = F_2(\bar{L}_2) - G; \quad \bar{q} = q(\bar{L}_1, \bar{L}_2); \\ &\bar{I}_1 + \bar{I}_2 = F_1(\bar{L}_1) - I_G, \quad I_1^d = \hat{I}_1^d(\bar{L}_1), \ I_2^d = \hat{I}_2^d(\bar{L}_1); \\ &\bar{I}_1 = bI_1^d, \quad \bar{I}_2 = bI_2^d, \quad \textit{with } b = (F_1(\bar{L}_1) - I_G)/(I_1^d + I_2^d); \\ &\bar{B}_1 = B_{1,0} + (p_1\bar{I}_1 - r_1(\bar{L}_1))/\bar{q}, \quad \bar{B}_2 = B_{2,0} + (p_1\bar{I}_2 - r_2(\bar{L}_2))/q. \end{aligned} \tag{25}$$

define a fixed-price equilibrium.

Proof. *By definition, $\bar{q} = q(\bar{L}_1, \bar{L}_2)$ is the equilibrium price on the bonds market, when levels of employment are $\bar{L}_1$ and $\bar{L}_2$.*

On the labor market, we have

$$\bar{L}_1 + \bar{L}_2 = \min\{L^0, L_1^d + L_2^d\}.$$

On the capital good market, we consider two cases:

$$\bar{L}_1 = L_1^d \qquad \text{and} \qquad \bar{L}_1 < L_1^d.$$

For $\bar{L}_1 = L_1^d$, the supply of capital good is $I^s = F_1(L_1^2)$.

If $I^s \leqslant F_1(L_1^3)$, sector 1 is not constrained, $L_1^3 \geqslant L_1^2$; then $\bar{L}_1 = L_1^2$ and $F_1(\bar{L}_1) = I^s$; we obtain

$$I^s \leqslant F_1(L_1^3) = I_1^d + I_2^d + I_G = I^d$$

and $\bar{I}_1 + \bar{I}_2 + I_G = \min\{I^s, I^d\} = I^s$.

If $I^s > F_1(L_1^3)$, then $L_1^3 < L_1^2$ and $\bar{L}_1 = L_1^3$; consequently

$$\bar{I}_1 + \bar{I}_2 + I_G = F_1(L_1^3) = \min\{I^s, I^d\} = I^d.$$

For $\bar{L}_1 < L_1^d$, sector 1 is constrained on the labor market, and $I^s = F_1(\bar{L}_1)$; consequently

$$\bar{I}_1 + \bar{I}_2 + I_G = I^s = \min\{I^d, I^s\}.$$

For the consumption good market, the proof is similar.

II. EFFECTS ON THE EQUILIBRIUM OF THE PARAMETERS

In this section we shall consider the effects, on the equilibrium, of variations in the parameters such as: public expenditure (consumption and investment), prices and wages.

II.1. *Framework of the Study.* We assume that current income only consists of wages:

$$D_1(L_1) = wL_1 \qquad \text{and} \qquad D_2(L_2) = wL_2$$

Given the multiplicity of types of equilibrium, we shall select three types for our study:

(A) Sector 1 with classical unemployment and sector 2 with Keynesian unemployment.

(B) Sector 1 with Keynesian unemployment ans sector 2 with classical unemployment.

(C) Both sectors with Keynesian unemployment.

In the selected types of equilibrium, the demands for labour in the sectors are satisfied and the demands for investment are as follows, according to the assumptions already made:

$$I_1^d = A_1(q(L_1, L_2), V) \quad \text{and} \quad I_2^d = A_2(q(L_1, L_2), V).$$

V stands for the other non-explicit parameters: prices, wages,.... Later, we shall be led to make use of assumptions about the investment functions, the demand of the sectors for money, the incidence of prices and wages on the expectations of households.

Assumptions about the investment functions:

$$\frac{\partial A_i}{\partial q} > 0, \qquad \frac{\partial A_i}{\partial w} > 0, \qquad \frac{\partial A_i}{\partial p_1} < 0 \qquad (i = 1, 2).$$

Such assumptions express the fact that the demand for investment in each sector is a decreasing function of the interest rate $(1/q)$ as well as a decreasing function of the relative capital–labour cost.

Assumptions about the demand of sectors for money:

$$\frac{\partial M_1}{\partial p_1} > 0, \qquad \frac{\partial M_2}{\partial p_2} > 0; \qquad \frac{\partial M_1}{\partial w} < 0, \qquad \frac{\partial M_2}{\partial w} < 0.$$

Money is, for the most part, destined for the payment of taxes and distribution of profits. All other things being equal, we may assume that they vary directly with the production price and inversely with the wage rate.

Assumptions about the expectations of households:

$$\frac{\partial P}{\partial w} > 0, \qquad \frac{\partial P}{\partial p_2} > 0.$$

The assumptions express the fact that expectations on nominal income are an increasing function of the nominal wage rate and of the consumer price.

II.2. *Results.* The details of the study are in the Appendix. We obtain the following results.

(A) *The case of classical unemployment in sector* 1 *and Keynesian unemployment in sector* 2.

Effect of public consumption G: An increase in G has no effect on the production of sector 1; government consumption has a stimulating effect on sector 2 *with the usual Keynesian multiplier* $1 + a/\beta$, *modified by possible effects on the expectation of households and firms liquid assets* (Appendix, relation (7).)

Effect of public investment I_G: An increase in I_G has also a positive multiplier effect on sector 2, *but such an effect is smaller than of public* consumption, for an equal level of expenditure (Appendix (8)). On sector 1, this policy reduces to the same degree rationed private investments, since the total poduction of capital goods does not depend on demand.

Effect of a change in the wage rate w (Appendix (9) and (10)): The increase in w has a stimulating effect on sector 2 through its effects on the income expectations of households, and a depressing effect on sector 1 through the decrease in profit.

Effects of a change in the price p_1 *of capital goods* (Appendix (11)): The increase in p_1 has a stimulating effect on sector 1 through the increase in its profit, and an effect on sector 2 made up of two terms: (1) the first is positive and corresponds to the additional creation of money by the government to pay for its investment; (2) the second is rather negative if we assume that the most important effect of p_1 is felt by the monetary reserves of sector 1 and that these reserves are increasing along p_1.

Consequently, the resulting effect on sector 2 may be positive or negative. It depends on whether the additional government expenditures resulting from the increase in the price p_1 exceed or do not exceed the additional monetary reserves decided upon by sector 1. In the medium run, this effect will in any case be positive on account of the increase in the income distributed by sector 1.

Changes in the price p_2 *of consumer goods* (Appendix (12) and (13)): These changes have no effect on the first sector. For the second sector, in addition to the usual depressing effect on the demand, there are effects of the expectations of households and of the decisions concerning monetary reserves made by the firms in sector 2. We may think that the first term is dominant, such that an increase in the consumer price exerts a depressing effect on sector 2.

(B) *The case of Keynesian unemployment in sector* 1 *and of classical unemployment in sector* 2.

Effect of public consumption G: An increase in G has no effect on sector 2 and a positive indirect effect on sector 1 (Appendix (17) and (18)).

Effect of public investment I_G: An increase in I_G has no effect on sector 2 and a direct multiplier positive effect on sector 1 (Appendix (19) and (20)).

Given the same increase in the public expenditures ($p_2 dG = p_1 dI_G$) it is obviously the investment expenditures which have the greater effect on the activity in sector 1. For public consumption expenditures, the increase in activity results from the effect on investment of a reduction in the interest rate ($dq > 0$). This investment can be financed by an increase in "unvoluntary savings" of households (whose consumption decreases by dG) enabling them to subscribe to additional securities issued by the industrial sectors.

Effect of the wage rate (Appendix (21)–(23)): A rise in nominal wages depresses the consumer goods sector through a reduction of the profitable production capacity. The effect is more questionable for the investment goods sector. Indeed, there is a positive effect owing to the relative decrease in capital cost; but, on the other hand, the interest rate rises. We may assume that the second effect prevails over the first and therefore, in the aggregate, a rise in nominal wages is in this case conducive to a reduction in activity and employment.

Effect of the variations of p_1 (Appendix (24)–(26)): An increase in capital cost lowers the investment demand. Furthermore, public expenditures for investments rises by $I_G dp_1$, and there is a multiplier effect similar to an increase in I_G. The resulting effect is indeterminate: the rate of interest decreases, the impact on the investment demand is thus the inverse of an increase in p_1 and, without a more precise specification of investment functions, the aggregate effect cannot be estimated.

Influence of the variations in p_2 (Appendix (27) and (28)): The rise in prices in the consumer goods sector increases the profitable production capacity in this sector and consequently increases employment. We may consider that the total investment demand increases, so that the activity in the investment goods sector also increases.

(C) *The case of Keynesian unemployment in both sectors.*

Effect of public consumption G: The effect on sector 2 is the usual Keynesian multiplier effect. The effect on sector 1 is a positive indirect effect which depends on the marginal profit rate in sector 2 (Appendix (31)–(33)). If we consider the aggregate effect of an increase dG on the volume of production and employment, the μ_G multiplier of aggregate production is defined by

$$p_1 dY_1 + p_2 dY_2 = \mu_G p_2 dG,$$

and the employment multiplier λ_G is defined by

$$dL_1 + dL_2 = \lambda_G p_2 dG.$$

The aggregate Keynesian multiplier μ_G is the sum of those multipliers obtained in the two previous cases, where one of the sectors is in classical unemployment (Appendix (34)).

We also find that as sector 2 gets closer to a situation of classical unemployment, the aggregate multipliers become lower.

Effect of public investment I_G: An increase in I_G has symmetric effects on the activity in the two sectors (Appendix (36) and (37)).

The aggregate multipliers of public investment μ_I and λ_I for production and employment are defined respectively by

$$p_1 dY_1 + p_2 dY_2 = \mu_I p_1 dI_G,$$
$$dL_1 + dL_2 = \lambda_I p_1 dI_G.$$

By comparing the public consumption and investment multiplier μ_G and μ_I, we see that any public expenditure policy designed to increase the volume of production, should privilege the sector where the marginal productivity of labour is the highest (Appendix (40)). And conversely a policy aiming at an increase in employment should privilege the sector with the lowest productivity (Appendix (41)). The two objectives thus appear contradictory.

Impact of the wage rate (Appendix (42)–(44)): An increase in the wage rate results in: (1) an increase in activity in the consumer good sector induced by income expectations of households; (2) a twofold effect in the investment goods sector; the first effect is the impact on investment which leads to a reduction of activity. The second effect, which is positive, is an indirect consequence of positive income expectations of households.

On the whole, the impact on activity and employment is therefore likely to be limited, though more favorable than in case 2, i.e., Keynesian equilibrium for sector 1, classical equilibrium for sector 2.

Variation in price p_1 of capital goods: In both sectors, an increase in p_1 induces opposite effects (see relations (45) and (46), Appendix) and it is difficult, without any further assumptions to reach a conclusion about the final results.

Variation in the price p_2 of consumer goods (Appendix (47) and (48)) The increase in the price of consumer goods has a negative effect on thei production, but for the same reasons as mentioned earlier, activity rises i the investment goods sector.

CONCLUSION

To recap the three cases under study:

	G	I_G	w	p_1	p_2
Sector 1 : classical	0	0	− −	+ +	0
Sector 2 : Keynesian	+ +	+	+	?	−
Sector 1 : Keynesian	+	+ +	−	?	+
Sector 2 : classical	0	0	− −	0	+ +
Sector 1 : Keynesian	+	+ +	?	?	+
Sector 2 : Keynesian	+ +	+	+	?	−

In the three cases examined, increases in public expenditures have a stimulating effect: in each case, at least one of the two sectors is in Keynesian unemployment. The effect is direct when it involves the sector in Keynesian unemployment. Otherwise it is indirect. Given equal government expenditures, the direct effect is more substantial than the indirect one.

The increase in the nominal wage rate results in the development of activity in the consumer goods sector when this sector is in Keynesian unemployment. Such an effect stems from the more favorable income expectations of households.

Conversely, the rise in nominal wages usually has a negative effect on the investment goods sector. The relative decrease in the cost of capital as compared to the cost of labour appears to be more than offset by the rise in the interest rate. Obviously, it always has a negative effect on a sector in classical unemployment.

The only clear cut effect of an increase in the price of investment goods is the stimulation of the corresponding sector when it is in classical unemployment. In the other cases, such an effect is not conclusive; it is made up of two contradictory effects: a decrease in the demand for investment which results from the rise in cost and a reflation through an increase in the actual value of public demand. In the consumer goods sector, we again find the same effects, but expressed indirectly.

A rise in the consumer price has a negative effect on the corresponding sector when it is Keynesian unemployment, and a positive effect when it is in classical unemployment. The impact on the investment goods sector is positive when this sector is in Keynesian unemployment.

The consequences of economic policies thus appear to be extremely different in effect depending on the sector to which they apply and on the nature of the equilibrium prevailing in the sector.

APPENDIX: MATHEMATICAL STUDY OF THE EFFECTS ON THE EQUILIBRIUM OF THE PARAMETERS

For the considered model, we have

$$M_c = p_2 G + p_1 I_G + M_{c,0} - \Delta M_E,$$

$$\begin{aligned} q(B_0 + Q) = {} & \frac{\gamma}{\beta}(P + p_2 G + M_{c,0}) + \left(\frac{\gamma}{\beta} + 1\right)(p_1 I_G - \Delta M_E) \\ & + p_2 F_2(L_2) - wL_1 - wL_2, \end{aligned} \tag{1}$$

$$C^d = \frac{\alpha}{p_2}\left[p_2(F_2(L_2) - G) + \left(\frac{\gamma}{\beta} + 1\right)(P + M_c)\right],$$

where $\Delta M_E = M_1 - M_{1,0} + M_2 - M_{2,0}$.

We deduce from (1) and $\alpha + \beta + \gamma = 1$

$$\begin{aligned} F_2(L_2) - C^d = {} & (1 - \alpha) F_2(L_2) \\ & - \frac{\alpha\gamma G}{\beta} - \frac{\alpha(1-\alpha)}{\beta p_2}(P + M_{c,0} - \Delta M_E + p_1 I_G), \end{aligned} \tag{2}$$

which is an increasing function of L_2; the demand for labour L_2^d of sector 2 is therefore defined by

$$\begin{aligned} & L_2^d = \min\{L_2^2, L_2^3\}, \\ & L_2^2 = F_2'^{-1}\left(\frac{w}{p_2}\right), \\ & L_2^3 = F_2^{-1}\left(\left(\frac{\alpha + \beta}{\beta}\right) G + \frac{\alpha}{\beta p_2}(P + M_{c,0} - \Delta M_E + p_1 I_G)\right). \end{aligned} \tag{3}$$

Similarly, with the assumptions made, we have for sector 1:

$$\begin{aligned} & L_1^d = \min\{L_1^2, L_1^3\}, \\ & L_1^2 = F_1'^{-1}\left(\frac{w}{p_1}\right), \\ & L_1^3 = F_1^{-1}(I_1^d + I_2^d + I_G). \end{aligned} \tag{4}$$

(A) *Study of the Case of Classical Unemployment in Sector* 1 *and Keynesian Unemployment in Sector* 2

This case is characterized by

$$
\begin{aligned}
F_1'(L_1) &= \frac{w}{p_1},\\
F_2(L_2) &= \left(1+\frac{\alpha}{\beta}\right)G+\frac{\alpha}{\beta p_2}(N+p_1 I_G),
\end{aligned}
\tag{5}
$$

where $N = P + M_{c,0} + M_{1,0} - M_1 + M_{2,0} - M_2$.

We thus obtain

$$
\begin{aligned}
F_1''(L_1)\,dL_1 &= \frac{1}{p_1}\,dw - \frac{w}{p_1^2}\,dp_1,\\
F_2'(L_2)\,dL_2 &= \left(1+\frac{\alpha}{\beta}\right)dG - \frac{\alpha}{\beta p_2^2}(N+p_1 I_G)\,dp_2\\
&\quad + \frac{\alpha}{\beta p_2}(dN + p_1\,dI_G + I_G\,dp_1).
\end{aligned}
\tag{6}
$$

(a) *Effect of public consumption G*:

$$
dY_2 = F_2'(L_2)\,dL_2 = \left(1+\frac{\alpha}{\beta}\right)dG + \frac{\alpha}{\beta p_2}\frac{\partial N}{\partial G}\,dG. \tag{7}
$$

(b) *Effect of public investment I_G*: Assuming that $\partial N/\partial I_G$ is negligible, we obtain

$$
dY_2 = F_2'(L_2)\,dL_2 = \frac{\alpha p_1}{\beta p_2}\,dI_G. \tag{8}
$$

Since $\alpha > \beta$, the effect of nominal public investment

$$
\frac{p_2\,dY_2}{p_1\,dI_G} = \frac{\alpha}{\beta} > 1
$$

is a multiplier effect, but such an effect in smaller than that of public consumption, for an equal level of expenditure

$$
\frac{p_2\,dY_2}{p_2\,dG} = 1 + \frac{\alpha}{\beta}.
$$

(c) *Effect of a change in the wage rate w*: Other prices being constant, we have

$$\begin{aligned} F''(L_1)\,dL_1 &= \frac{1}{p_1}\,dw, \\ F_2'(L_2)\,dL_2 &= \frac{\alpha}{\beta p_2}\frac{\partial N}{\partial w}\,dw, \end{aligned} \tag{9}$$

$$\frac{\partial N}{\partial w} = \frac{\partial P}{\partial w} - \frac{\partial M_1}{\partial w} - \frac{\partial M_2}{\partial w}. \tag{10}$$

The effects of an increase in w are

$$\frac{dY_2}{dw} = \frac{\alpha}{\beta p_2}\frac{\partial N}{\partial w} > 0,$$

$$\frac{dY_1}{dw} = \frac{F_1'(L_1)}{p_1 F_1''(L_1)} < 0.$$

(d) *Effects of a change in the price p_1 of capital goods*:

$$\begin{aligned} F_1''(L_1)\,dL_1 &= -\frac{w}{p_1^2}\,dp_1, \\ F_2'(L_2)\,dL_2 &= \frac{\alpha}{\beta p_2}\left(\frac{\partial N}{\partial p_1} + I_G\right)dp_1. \end{aligned} \tag{11}$$

The increase in p_1 has

a stimulating effect on sector 1 through the increase in its profit:

$$\frac{dY_1}{dp_1} = -\frac{w}{p_1^2}\frac{F_1'(L_1)}{F_1''(L_1)} > 0;$$

an effect on sector 2 made up of two terms:

(1) $$\frac{\alpha}{\beta p_2} I_G > 0$$

(2) $$\frac{\alpha}{\beta p_2}\frac{\partial N}{\partial p_1} \simeq -\frac{\alpha}{\beta p_2}\frac{\partial M_1}{\partial p_1};$$

assuming hat the most important effect of p_1 is felt by the monetary reserves of sector 1 and that these reserves are increasing along p_1, we obtain a negative effect.

(e) *Changes in the price p_2 of consumer goods*: These changes have no effect on the first sector. For the second sector, we have

$$F_2'(L_2)\,dL_2 = -\frac{\alpha}{\beta p_2^2}(N + p_1 I_G)\,dp_2 + \frac{\alpha}{\beta p_2}\frac{\partial N}{\partial p_2}\,dp_2, \tag{12}$$

$$\frac{\partial N}{\partial p_2} \simeq \frac{\partial P}{\partial p_2} - \frac{\partial M_2}{\partial p_2}. \tag{13}$$

(B) *Study of the Case of Keynesian Unemployment in Sector* 1 *and of Classical Unemployment in Sector* 2

This case is characterized by

$$\begin{aligned} F_1(L_1) &= I_G + A_1(q, V) + A_2(q, V) = I_G + A(q, V), \\ F_2'(L_2) &= \frac{w}{p_2}. \end{aligned} \tag{14}$$

We thus obtain

$$\begin{aligned} dY_1 = F_1'(L_1)\,dL_1 &= dI_G + \frac{\partial A}{\partial q}\,dq + \frac{\partial A}{\partial V}\,dV, \\ F_2''(L_2)\,dL_2 &= \frac{1}{p_2}\,dw - \frac{w}{p_2^2}\,dp_2. \end{aligned} \tag{15}$$

The variation in q is derived from (1):

$$\begin{aligned} &(B_0 + Q)\,dq + q\,dQ \\ &\quad = \frac{\gamma}{\beta}(dP + p_2 dG + G dp_2) + \left(\frac{\gamma}{\beta} + 1\right)(p_1 dI_G + I_G dp_1) \\ &\quad\quad + p_2 F_2'(L_2)\,dL_2 + F_2(L_2)\,dp_2 - w(dL_1 + dL_2) - (L_1 + L_2)\,dw. \end{aligned} \tag{16}$$

(a) *Effect of public consumption G*:

$$dY_1 = \frac{\partial A}{\partial q}\frac{1}{(B_0 + Q)}\left(\frac{\alpha}{\beta}\,p_2 dG - w\,dL_1\right). \tag{17}$$

Let

$$k = \frac{A}{q}\frac{1}{B_0 + Q} \text{ et } \frac{1}{\delta} = 1 + \frac{kw}{F_1'(L_1)}.$$

Under the assumptions, k is **positive and we therefore have** $0 < \delta < 1$, such that (17) becomes

$$dY_1 = k\delta \frac{\gamma}{\beta} p_2 dG. \tag{18}$$

(b) *Effect of public investment* I_G:

$$dY_1 = dI_G + \frac{\partial A}{\partial q} \frac{1}{B_0 + Q} \left(\left(\frac{\gamma}{\beta} + 1 \right) p_1 dI_G - w dL_1 \right) \tag{19}$$

and we obtain

$$dY_1 = \delta \left(1 + k \left(\frac{\gamma}{\beta} + 1 \right) p_1 \right) dI_G . \tag{20}$$

(c) *Effect of the wage rate*: Other prices being constant, we have

$$\begin{aligned} F''(L_2)\, dL_2 &= \frac{1}{p_2} dw, \\ F_1'(L_1)\, dL_1 &= \frac{\partial A}{\partial w} dw + \frac{\partial A}{\partial q} \frac{1}{B_0 + Q} [p_2 F_2'(L_2)\, dL_2 \\ &\quad - w(dL_1 + dL_2) - (L_1 + L_2)\, dw]. \end{aligned} \tag{21}$$

We obtain, with $p_2 F_2'(L_2) = w$,

$$\begin{aligned} dY_1 &= F_1'(L_1)\, dL_1 = \delta \left[\frac{\partial A}{\partial w} - k(L_1 + L_2) \right] dw, \\ dY_2 &= \frac{F_2'(L_2)}{p_2 F_2''(L_2)} dw. \end{aligned} \tag{22}$$

By substituting the corresponding terms from (22) in the expression of dq for the variation dL_1, we obtain

$$dq = \frac{-\delta}{B_0 + Q} \left[\frac{w}{F_1'(L_1)} \frac{\partial A}{\partial w} + (L_1 + L_2) \right] dw. \tag{23}$$

(d) *Effect of the variations of* p_1:

$$\begin{aligned} dY_2 &= 0, \\ F_1'(L_1)\, dL_1 &= \frac{\partial A}{\partial p_1} dp_1 + k \left[\left(\frac{\gamma}{\beta} + 1 \right) I_G dp_1 - w dL_1 \right], \end{aligned} \tag{24}$$

$$dY_1 = \delta \left[\frac{\partial A}{\partial p_1} + k \left(\frac{\gamma}{\beta} + 1 \right) I_G \right] dp_1 . \tag{25}$$

Substituting the corresponding terms from (25) in the expression of dq for the variation of L_1, we obtain

$$dq = \frac{\delta}{B_0 + Q}\left[\left(\frac{\gamma}{\beta} + 1\right) I_G - \frac{w}{F_1'}\frac{\partial A}{\partial p_1}\right] dp_1. \tag{26}$$

(e) *Influence of the variations in p_2:*

$$\begin{aligned} F_2''(L_2)\,dL_2 &= -\frac{w}{p_2^2}\,dp_2, \\ F_1'(L_1)\,dL_1 &= \frac{\partial A}{\partial p_2}\,dp_2 + k\left[\frac{\gamma}{\beta} G dp_2 - w dL_1\right], \end{aligned} \tag{27}$$

$$\begin{aligned} dY_1 &= \delta\left[\frac{\partial A}{\partial p_2} + k\frac{\gamma}{\beta} G dp_2\right], \\ dY_2 &= -\frac{wF_2'(L_2)}{p_2^2 F_2''(L_2)}\,dp_2 = -\frac{w^2}{p_2^3 F_2''(L_2)}\,dp_2. \end{aligned} \tag{28}$$

(C) *Study of the Case of Keynesian Unemployment in Both Sectors*

This case is defined by

$$\begin{aligned} F_1(L_1) &= I_G + A_1(q, V) + A_2(q, V) = I_G + A(q, V), \\ F_2(L_2) &= \frac{\alpha + \beta}{\beta} G + \frac{\alpha}{\beta p_2}(N + p_1 I_G), \end{aligned} \tag{29}$$

with $N = P + M_{c,0} + M_{1,0} - M_1 + M_{2,0} - M_2$; we thus obtain

$$\begin{aligned} F_1'(L_1)\,dL_1 &= dI_G + \frac{\partial A}{\partial q}\,dq + \frac{\partial A}{\partial V}\,dV, \\ F_2'(L_2)\,dL_2 &= \frac{\alpha + \beta}{\beta}\,dG + \frac{\alpha}{\beta p_2}\left(\frac{\partial N}{\partial V}\,dV + p_1 dI_G + I_G dp_1\right) \\ &\quad - \frac{\alpha}{\beta p_2^2}(N + p_1 I_G)\,dp_2. \end{aligned} \tag{30}$$

The variation of q is derived from (1)

$$\begin{aligned} &(B_0 + Q)\,dq + q\,dQ \\ &\quad = \frac{\gamma}{\beta}(dP + p_2 dG + G dp_2) + \left(\frac{\gamma}{\beta} + 1\right)(p_1 dI_G + I_G dp_1) \\ &\qquad + p_2 F_2'(L_2)\,dL_2 + F_2(L_2)\,dp_2 - w(dL_1 + dL_2) - (L_1 + L_2)\,dw. \end{aligned}$$

The effect of a public consumption and investment variation on N
considered to be negligible.

(a) *Effect of public consumption G:* The effect on sector 2 is express
by

$$dY_2 = F_2'(L_2)\,dL_2 = \left(1 + \frac{\alpha}{\beta}\right) dG. \qquad (3$$

$1 + \alpha/\beta$ is the usual Keynesian multiplier.

The effect on sector 1 is defined by

$$dY_1 = \frac{\partial A}{\partial q}\frac{1}{B_0 + Q}\left[\frac{\gamma}{\beta} p_2\,dG - w\,dL_1 + (p_2 F_2'(L_2) - w)\,dL_2\right]. \qquad (32)$$

Let

$$k = \frac{1}{B_0 + Q}\frac{\partial A}{\partial q},$$

$$\frac{1}{\delta} = 1 + k\frac{w}{F_1'(L_1)},$$

$$\sigma_i = \frac{p_i F_i'(L_i) - w}{p_i F_i'(L_i)} \qquad (i = 1, 2).$$

σ_i is the marginal profit rate in sector i.

We obtain

$$dY_1 = F_1'(L_1)\,dL_1 = k\delta p_2\left[\frac{\gamma}{\beta} + \left(1 + \frac{\alpha}{\beta}\right)\sigma_2\right] dG. \qquad (33)$$

Let us now examine the aggregate effect of an increase dG on the volume of production and employment.

The μ_G multiplier of aggregate production is defined by

$$p_1\,dY_1 + p_2\,dY_2 = \mu_G\,p_2\,dG,$$

$$\mu_G = \frac{dY_2}{dG} + \frac{p_1}{p_2}\frac{dY_1}{dG} = 1 + \frac{\alpha}{\beta} + k\delta p_1\left(\frac{\gamma}{\beta} + \left(1 + \frac{\alpha}{\beta}\right)\sigma_2\right). \qquad (34)$$

The employment multiplier λ_G is defined by

$$dL_1 + dL_2 = \lambda_G\,p_2\,dG,$$

$$\lambda_G = \frac{1}{p_2}\frac{dL_1}{dG} + \frac{1}{p_2}\frac{dL_2}{dG},$$

$$\lambda_G = \frac{1}{p_2 F_2'(L_2)}\left(1 + \frac{\alpha}{\beta}\right) + \frac{k\delta p_1}{p_1 F_1'(L_1)}\left(\frac{\gamma}{\beta} + \left(1 + \frac{\alpha}{\beta}\right)\sigma_2\right). \qquad (35)$$

(b) *Effect of public investment I_G*: We obtain for sector 2

$$dY_2 = F'_2(L_2)\, dL_2 = \frac{\alpha p_1}{\beta p_2}\, dI_G, \tag{36}$$

and finally for sector 1

$$dY_1 = F'_1(L_1)\, dL_1 = \delta \left[1 + kp_1 \left(1 + \frac{\gamma}{\beta} + \frac{\alpha}{\beta}\sigma_2\right)\right] dI_G. \tag{37}$$

The aggregate multipliers of public investment μ_I and λ_I are

$$\begin{aligned} \mu_I &= \frac{dY_1}{dI_G} + \frac{p_2}{p_1}\frac{dY_2}{dI_G}, \\ \mu_I &= \frac{\alpha}{\beta} + \delta + kp_1\delta\left(1 + \frac{\gamma}{\beta} + \frac{\alpha}{\beta}\sigma_2\right), \\ \lambda_I &= \frac{1}{p_1 F'_1(L_1)}\frac{dY_1}{dI_G} + \frac{1}{p_1 F'_2(L_2)}\frac{dY_2}{dI_G}, \end{aligned} \tag{38}$$

$$\lambda_I = \frac{\delta}{p_1 F'_1(L_1)}\left(1 + kp_1\left(1 + \frac{\gamma}{\beta} + \frac{\alpha}{\beta}\sigma_2\right)\right) + \frac{\alpha}{\beta p_2 F'_2(L_2)}. \tag{39}$$

We obtain

$$\begin{aligned} \mu_G - \mu_I &= 1 + \frac{\alpha}{\beta} + k\delta p_1\left(\frac{\gamma}{\beta} + \left(1 + \frac{\alpha}{\beta}\right)\sigma_2\right) \\ &\quad - \left[\frac{\alpha}{\beta} + \delta + k\delta p_1\left(1 + \frac{\gamma}{\beta} + \frac{\alpha}{\beta}\sigma_2\right)\right] \\ &= k\delta p_1 [\sigma_2 - \sigma_1]. \end{aligned} \tag{40}$$

From this we derive

$$\mu_G \gtrless \mu_I \Leftrightarrow \sigma_2 \gtrless \sigma_1, \qquad \text{i.e.} \qquad \frac{p_2F'_2 - w}{p_2F'_2} \gtrless \frac{p_1F'_1 - w}{p_1F'_1},$$

and we see that $\sigma_2 \gtrless \sigma_1 \Leftrightarrow p_1F'_1 \lessgtr p_2F'_2$.

Conversely, if we consider the employment multipliers λ_G and λ_I, we obtain

$$\begin{aligned} \lambda_G - \lambda_I &= \delta\left[\frac{1}{p_2F'_2} - \frac{1}{p_1F'_1}\right], \\ \lambda_G &\gtrless \lambda_I \Leftrightarrow p_1F'_1 \gtrless p_2F'_2. \end{aligned} \tag{41}$$

(c) *Impact of the wage rate*: Other prices being constant, we have

$$\begin{aligned} F_2'(L_2)\,dL_2 &= \frac{\alpha}{\beta p_2}\frac{\partial N}{\partial w}\,dw, \\ F_1'(L_1)\,dL_1 &= \frac{\partial A}{\partial q}\,dq + \frac{\partial A}{\partial w}\,dw \end{aligned} \tag{42}$$

$$\begin{aligned} dq = \frac{1}{B_0 + Q}\Bigg[&\frac{\gamma}{\beta}\frac{\partial P}{\partial w}\,dw + (p_2F_2'(L_2) - w)\,dL_2 - w\,dL_1 \\ &- (L_1 + L_2)\,dw\Bigg]. \end{aligned} \tag{43}$$

By using (43), we have

$$\begin{aligned} dY_1 = F_1'(L_1)\,dL_1 &= \delta\left[\frac{\partial A}{\partial w} - k(L_1 + L_2)\right]dw \\ &+ k\delta\left[\frac{\gamma}{\beta}\frac{\partial P}{\partial w} + \frac{\alpha}{\beta}\frac{\partial N}{\partial w}\sigma_2\right]dw. \end{aligned} \tag{44}$$

According to the hypotheses, $\partial N/\partial w$ is positive but undoubtedly rather low; $\partial A/\partial w$ is positive (capital–labour subsitution effect).

(d) *Variation in price p_1 of capital goods*:

$$\begin{aligned} F_1'(L_1)\,dL_1 &= \frac{\partial A}{\partial p_1}\,dp_1 + \frac{\partial A}{\partial q}\,dq, \\ F_2'(L_2)\,dL_2 &= \frac{\alpha}{\beta p_2}\left(\frac{\partial N}{\partial p_1}\,dp_1 + I_G\,dp_1\right), \\ dq &= \frac{1}{B_0 + Q}\left[\frac{\gamma}{\beta}\frac{\partial P}{\partial p_1}\,dp_1 + \left(\frac{\gamma}{\beta} + 1\right)I_G\,dp_1\right. \\ &\qquad \left. + (p_2F_2'(L_2) - w)\,dL_2 - w\,dL_1\right]. \end{aligned} \tag{45}$$

By substituting expression dq, we finally obtain

$$\begin{aligned} dY_1 = F_1'(L_1)\,dL_1 &= \delta\frac{\partial A}{\partial p_1}\,dp_1 \\ &+ k\delta\left[\frac{\gamma}{\beta}\frac{\partial P}{\partial p_1} + \frac{\alpha}{\beta}\frac{\partial N}{\partial p_1}\sigma_2 + \left(1 + \frac{\gamma}{\beta} + \frac{\alpha\sigma_2}{\beta}\right)I_G\right]dp_1. \end{aligned} \tag{46}$$

The increase in p_1 has a positive impact $(I_G dp_1)$ and a negative impact $(\partial N/\partial p_1 \simeq -\partial M_1/\partial p_1)$ on the consumer goods sector.

The effect on sector 1 is of a similar nature to the effect studied in B: $\partial A/\partial p_1$ is negative, but the second term, which represents a muliplier effect of public expenditure, is positive. Here, too, the resulting effect is indeterminate.

(e) *Variation in the price p_2 of consumer goods*: The effects are defined by the following relations:

$$\begin{aligned}
&(1)\ F_1'(L_1)\,dL_1 = \frac{\partial A}{\partial p_2} + \frac{\partial A}{\partial q}\,dq,\\
&(2)\ F_2'(L_2)\,dL_2 = \frac{\alpha}{\beta p_2}\left(\frac{\partial N}{\partial p_2} - \frac{1}{p_2}(N + p_1 I_G)\right) dp_2, \qquad (47)\\
&(3)\ dq = \frac{1}{B_0 + Q}\left[\frac{\gamma}{\beta}\frac{\partial P}{\partial p_2}dp_2 + \frac{\gamma}{\beta} G\,dp_2 + (p_2 F_2'(L_2) - w)\,dL_2\right.\\
&\qquad\qquad \left. + F_2'(L_2)\,dp_2 - w\,dL_1\right].
\end{aligned}$$

Substituting dq and taking into account expression (29) of $F_2(L_2)$, we finally obtain

$$\begin{aligned}
F_1'(L_1)\,dL_1 = \delta \frac{\partial A}{\partial p_2}dp_2 + \delta k &\left[\frac{\alpha}{\beta}\frac{\partial P}{\partial p_2} + \left(\frac{\gamma}{\beta} + \left(\frac{\alpha}{\beta} + 1\right)\sigma_2\right)\right.\\
&\left. + \frac{\alpha}{\beta}\sigma_2\frac{\partial N}{\partial p_2} + (1 - \sigma_2)F_2(L_2)\right] dp_2. \qquad (48)
\end{aligned}$$

ACKNOWLEDGMENTS

The authors are grateful to J. M. Grandmont and P. A. Muet, and other colleagues at CEPREMAP and to a referee of this Journal for their helpful suggestions.

REFERENCES

1. R. J. Barro and H. I. Grossman, General disequilibrium model of income and employment, *Amer. Econ. Rev.* **61** (1971), 82–93.
2. R. J. Barro and H. I. Grossman, "Money, Employment and Inflation," Cambridge Univ. Press, Cambridge, 1976.
3. J. P. Benassy, Neo-keynesian disequilibrium in a monetary economy, *Rev. Econ. Studies* **42** (1975), 503–523.

4. J. P. BENASSY, *A neokeynesian model of price and quantity determination in desequilibrium*, *in* "Equilibrium and Disequilibrium in Economic Theory" (G. Schwodiauer, Ed.), pp. 511–544, Dordrecht, 1977. (Proceedings of a conference held in Vienna, Austria, 1974.)
5. R. W. CLOWER, The Keynesian counterrevolution: A theoretical appraisal, *in* "The Theory of Interest Rates" (F. H. Hahn and Brechling, Eds.), Mcmillan, London, 1965.
6. J. DREZE, Existence of an equilibrium under price rigidity and quantity rationing, *Internat. Econ. Rev.* **16** (1975), 301–320.
7. J.M. GRANDMONT, Temporary general equilibrium theory, *Econometrica* **45** (1977), 535–572.
8. J.M. GRANDMONT AND G. LAROQUE, On Keynesian temporary equilibria, *Rev. Econ. Studies* **43** (1976), 53–67.
9. K. HILDENBRAND AND W. HILDENBRAND, On Keynesian equilibria with unemployment and quantity rationing, *J. Econ. Theory* **18** (1978), 255–277.
10. B. HOOL, "Monetary and Fiscal Policies in Equilibria with Rationing," SSRI, University of Wisconsin, Madison, 1977.
11. E. MALINVAUD, "The Theory of Unemployment Reconsidered," Blackwell, Oxford, 1977.
12. E. MALINVAUD AND Y. YOUNES, Une Nouvelle Formulation Générale pour l'Étude des Fondements Microéconomiques de la Macroéconomie, *Cahiers Sém. d'Économétrie C.N.R.S.* **18** (1977), 63–112.
13. J. MUELLBAUER AND R. PORTES, Macroeconomic models with quantity rationing, *Econ. J.* **88** (1978), 788–821.
14. Y. YOUNES, On the role of money in the process of exchange and the existence of a non-Walrasian equilibrium, *Rev. Econ. Studies* **42** (1975), 489–501.

Sufficiency of the Pontryagin Conditions for Optimal Control when the Time Horizon Is Free*

ARTHUR J. ROBSON

Department of Economics, University of Western Ontario,
London N6A 5C2, Canada

Received July 28, 1978; revised January 16, 1980

1. INTRODUCTION

The Maximum Principle of optimal control theory states the necessary conditions which a solution must satisfy (see Pontryagin, *et al.* [6], or Lee and Markus [3]). It was shown by Mangasarian [4] that these conditions are also sufficient when the Hamiltonian is concave in the control and state variables. However, Mangasarian explicitly considers only the case with a fixed time interval. Arrow and Kurz [1] offer a generalization of the Mangasarian result, requiring only that the maximized Hamiltonian be conave in the state variables. (See Seierstad and Sydsaeter [7] for a rigorous proof of this result.) Again, the time interval is taken to be fixed. There are, however, a number of applications of optimal control in economics where the final "time," say, is freely chosen. For example, the overall radius of a city may be chosen in models of urban location. (See Mirrlees [5].) Koopmans [2] considers a model of an optimal apocalypse with depletable resources. (Koopmans does not use the Maximum Principle but offers a direct sufficiency proof for the particular case considered.) It is the purpose of this note to show that, given an intuitively appealing condition in addition to concavity of the maximized Hamiltonian, sufficiency still obtains. The proof proceeds by reducing the variable time case by a simple "trick" to the known fixed time situation.

---

* This was written while the author was a Visiting Fellow at the Australian National University.

0022-0531/81/030437-09$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

2. PROBLEM FORMULATION

The basic problem is taken to be to maximize the objective functional

$$J = \int_{t_0}^{t_1} f^0(x, u, t)\, dt$$

subject to the "equations of motion" of the state variables

$$\dot{x}_i = f^i(x, u, t), \qquad i = 1,\dots, n,$$

where these functions and the partial derivatives

$$\frac{\partial f^i}{\partial x_j}, \frac{\partial f^i}{\partial t}, \qquad i = 0,\dots, n; \quad j = 1,\dots, n,$$

are continuous. The control vector $u(t)$ satisfies for all $t \in [t_0, t_1]$,

$$u(t) \in \Omega \subset R^m,$$

where Ω is a time-independent control set. "Admissible" controls are taken to be piecewise continuous, although the result also holds for bounded measurable controls. The boundary conditions for the state variables are taken to be

$$\begin{aligned} &x_i(t_0) = x_i^0, && i = 1,\dots, n,\\ &x_i(t_1) = x_i^f, && i = 1,\dots, p,\\ &x_i(t_1) \text{ freely chosen}, && i = p + 1,\dots, n. \end{aligned}$$

More general boundary conditions may be treated, with, say, an initial r_0-dimensional smooth manifold, S_0 and a final r_1-dimensional smooth manifold, S_1. In order for sufficiency to be proven, additional hypotheses are needed concerning the convexity of S_0 and S_1. (For fixed time and time optimal sufficiency theorems with such terminal manifolds, see Lee and Markus [3, pp. 340–360].) Finally, although the initial time t_0 is taken to be fixed, the terminal time, t_1, is freely chosen.

As is shown in Pontryagin [6], in order that $\{u^*(t)\}$, $\{x^*(t)\}$, and t_1^* be an optimal admissible control, associated path of the state variables, and terminal time, respectively, it is necessary that there exist "costate variables" $\{\psi_i(t)\}$ $i = 0,\dots, n$, not ever all zero, ψ_0 a non-negative constant, such that the Hamiltonian,

$$\mathscr{H}(x, u, \psi, t) = \sum_{i=0}^{n} \psi_i f^i(x, u, t)$$

is maximized by $u^*(t)$ over choice of $u \in \Omega$. The costate variables satisfy the adjoint equations

$$\dot{\psi}_i = -\frac{\partial \mathscr{H}}{\partial x_i}, \qquad i = 1, \ldots, n,$$

with

$$\psi_i(t_1^*) = 0, \qquad i = p + 1, \ldots, n,$$

corresponding to free choice of the associated state variables. Finally, corresponding to the free choice of terminal time it is necessary that

$$\mathscr{H}(x^*, u^*, \psi, t_1^*) = 0.$$

3. SUFFICIENCY THEOREM

Suppose that the maximized Hamiltonian

$$\begin{aligned} M(x, \psi, t) &= \operatorname*{Max}_{u \in \Omega} \mathscr{H}(x, u, \psi, t) \\ &= \operatorname*{Max}_{u \in \Omega} \sum_{i=0}^{n} \psi_i f^i(x, u, t) \end{aligned}$$

satisfies the following assumptions:

A.1 $\psi_0 > 0$ *and thus* $\psi_0 = 1$ *without loss of generality;*

A.2 $M(x, \psi, t)$ *is concave in* x;

A.3
$$\begin{aligned} M(x^*(t), \psi(t), t) &\geqslant 0, \qquad t \in [t_0, t_1^*], \\ M(x^*(t_1^*), \psi(t_1^*), t) &\leqslant 0, \qquad t \in (t_1^*, \infty), \end{aligned}$$

then the above Pontryagin necessary conditions are also sufficient.

NOTE. Assumption A.3 can be guaranteed if

$$\frac{\partial \mathscr{H}}{\partial t}(x, u, \psi, t) \leqslant 0,$$

which holds if, for example,

$$\frac{\partial f^0}{\partial t}(x, u, t) \leqslant 0 \qquad \text{and} \qquad \frac{\partial f^i}{\partial t}(x, u, t) = 0, \qquad i = 1, \ldots, n.$$

For then, the first part follows from

$$\frac{dM}{dt}(x^*, \psi, t) = \frac{\partial \mathscr{H}}{\partial t}(x^*, u^*, \psi, t) \leqslant 0, \qquad t \in [t_0, t_1^*].$$

Also,

$$\begin{aligned} \mathscr{H}(x^*(t_1^*), u, \psi(t_1^*), t) &\leqslant \mathscr{H}(x^*(t_1^*), u, \psi(t_1^*), t_1^*) \\ &\leqslant M(x^*(t_1^*), \psi(t_1^*), t_1^*) = 0 \end{aligned}$$

for all $u \in \Omega$, $t > t_1^*$, so that the second part follows.

4. Proof

Consider the following new problem with fixed terminal time T. The objective functional is

$$\bar{J} = \int_{t_0}^{T} v f^0(x, u, t)\, dt,$$

which is to be maximized subject to the equations of motion

$$\dot{x}_i = v f^i(x, u, t), \qquad i = 1, \dots, n,$$

over choice of controls

$$(u, v) \in \Omega \times [0, 1] \subset R^{m+1}.$$

Note then that v has been introduced as a new control variable. It acts as a "switch" for the entire problem, enabling a fixed terminal time to be formally considered. The boundary conditions are now taken to be

$$\begin{aligned} &x_i(t_0) = x_i^0, && i = 1, \dots, n, \\ &x_i(T) = x_i^f, && i = 1, \dots, p, \\ &x_i(T) \text{ freely chosen}, && i = p + 1, \dots, n. \end{aligned}$$

Necessary conditions for $\{\bar{u}^*(t), \bar{v}^*(t)\}$ and $\{\bar{x}^*(t)\}$ to be an optimal admissible control and associated path of the state variables, are that there exist costate variables $\{\bar{\psi}_i(t)\}$ $i = 0, \dots, n$, not ever all zero, $\bar{\psi}_0$ a non-negative constant, such that

$$\begin{aligned} \mathscr{H}(\bar{x}^*, u, v, \bar{\psi}, t) &= v \sum_{i=0}^{n} \bar{\psi}_i f^i(\bar{x}^*, u, t) \\ &= v \mathscr{H}(\bar{x}^*, u, \bar{\psi}, t) \end{aligned}$$

is maximized by $(\bar{u}^*, \bar{v}^*)$ in $\Omega \times [0, 1]$. The adjoint equations are

$$\dot{\bar{\psi}}_i = -\frac{\partial \bar{\mathscr{H}}}{\partial x_i}, \qquad i = 1, \dots, n,$$

with

$$\bar{\psi}_i(T) = 0, \qquad i = p + 1, \dots, n.$$

Consider now any solution of the necessary conditions of the original problem, given the assumptions of the sufficiency theorem. Suppose that $\{u^*(t)\}$, $\{x^*(t)\}$, t_1^*, and $\{\psi(t)\}$ are the admissible control, state variables, terminal time, and costate variables, respectively. It is desired to demonstrate that these characterize an optimum. Now extend the above functions to any larger interval $[t_0, T]$ as follows

$$\begin{aligned}
\bar{u}^*(t) &= u^*(t), && t \in [t_0, t_1^*], \\
&= u^*(t_1^*), && t \in (t_1^*, T], \\
\bar{v}^* &= 1, && t \in [t_0, t_1^*], \\
&= 0, && t \in (t_1^*, T], \\
\bar{x}^*(t) &= x^*(t), && t \in [t_0, t_1^*], \\
&= x^*(t_1^*), && t \in (t_1^*, T], \\
\bar{\psi}(t) &= \psi(t), && t \in [t_0, t_1^*], \\
&= \psi(t_1^*) && t \in (t_1^*, T].
\end{aligned}$$

Then $\{\bar{u}^*(t), \bar{v}^*(t)\}$, $\{\bar{x}^*(t)\}$, and $\{\bar{\psi}(t)\}$ satisfy the necessary conditions for the new problem for any $T \geqslant t_1^*$. Indeed, consider firstly $t \in [t_0, t_1^*]$. Since $v \geqslant 0$, $\bar{u}^*(t) \in \Omega$ maximizes $\bar{\mathscr{H}}(\bar{x}^*, u, v, \bar{\psi}, t)$ simply because $u^*(t) \in \Omega$ maximizes $\mathscr{H}(x^*, u, \psi, t)$. Then it follows from A.3 that $\bar{v}^*(t) \in [0, 1]$ maximizes $\bar{\mathscr{H}}(\bar{x}^*, \bar{u}^*, v, \bar{\psi}, t)$ for all $t \in [t_0, T]$. Then, for $t \in (t_1^*, T]$ the choice of $\bar{u}^*(t) \in \Omega$ is irrelevant so that $\bar{u}^*(t)$ maximizes $\bar{\mathscr{H}}(\bar{x}^*, u, v, \bar{\psi}, t)$ for all $t \in [t_0, T]$. The new variables clearly satisfy the appropriate boundary conditions as spelled out for the new problem. Finally, it is readily seen that the costate variables $\bar{\psi}_i$, $i = 0, \dots, n$, satisfy the appropriate adjoint equations and boundary conditions for the new problem, and by A.1, $\bar{\psi}_0 \equiv 1$.

Now the maximized value of the new Hamiltonian is, say

$$\begin{aligned}
\bar{M}(x, \bar{\psi}, t) &= M(x, \psi, t), && t \in [t_0, t_1^*], \\
&= 0, && t \in (t_1^*, T].
\end{aligned}$$

which is concave in x by A.2. Hence the usual sufficiency theorem may be invoked (as in Arrow and Kurz [1]) to conclude that $\{\bar{u}^*(t), \bar{v}^*(t)\}$ and $\{\bar{x}^*(t)\}$ are optimal for the new problem. Suppose now that $\{u(t)\}$, $\{x(t)\}$, and t_1 are any other admissible control, associated state variables, and terminal time for the original problem. Then, for any $T \geqslant \max[t_1^*, t_1]$ it must be true that

$$\int_{t_0}^{T} \bar{v} f^0(\bar{x}, \bar{u}, t)\, dt \leqslant \int_{t_0}^{T} \bar{v}^* f^0(\bar{x}^*, \bar{u}^*, t)\, dt,$$

where

$$\begin{aligned} \bar{v}(t) &= 1, \qquad t \in [t_0, t_1], \\ &= 0, \qquad t \in (t_1, T], \end{aligned}$$

and $\{\bar{x}(t)\}$ and $\{\bar{u}(t)\}$ extend $\{x(t)\}$ and $\{u(t)\}$ to $[t_0, T]$ precisely as $\{\bar{x}^*(t)\}$ and $\{\bar{u}^*(t)\}$ extend $\{x^*(t)\}$ and $\{u^*(t)\}$. Thus

$$\int_{t_0}^{t_1} f^0(x, u, t)\, dt \leqslant \int_{t_0}^{t_1^*} f^0(x^*, u^*, t)\, dt,$$

so that the necessary conditions of the original problem are also sufficient, as was to be proved.

5. An Application

As an example of the use of this theorem consider the following optimal urban model. (This is a simple version of Mirrlees [5].) Suppose that there are N identical individuals, each with utility

$$u(s, z), \qquad u_s(s, z) > 0,\ u_z(s, z) > 0,$$

where s is space and z is a composite consumption good. These individuals are to be located within a ring-shaped residential region with fixed inner radius, I, and variable outer radius, R. Suppose that a fraction, θ, of arc is devoted to residences. (The remainder is perhaps used for roads. However, for the sake of simplicity, the mechanism by which land used for roads commands rent will not be specified.) Thus

$$n(x)\, s(x) = 2\pi\theta x,$$

where $n(x)\,dx$ is the number of individuals located between x and $x + dx$. Also, then

$$N = \int_l^R n(x)\,dx.$$

Suppose that there is available a fixed amount of the composite commodity, Y, say. This is used for individual allocations of the composition good, z, but also to rent land at a fixed rent $\bar{r}$ (corresponding perhaps to agricultural use), and to pay for transportation, which is assumed to cost $t \cdot x$ for each individual living at x. (Note that, since the rent commanded by land in transportation is not specified, the relevant margin is between the residential use of land and its use in agriculture. Hence, only residential land is explicitly rented at rate $\bar{r}$.) Thus

$$Y = \int_l^R \{n(x)\,z(x) + 2\pi\theta x\bar{r} + n(x)\,tx\}\,dx.$$

A Benthamite welfate criterion is taken to be

$$W = \int_l^R u(s(x), z(x))\,dx = \int_l^R u\left(\frac{2\pi\theta x}{n(x)}, z(x)\right) n(x)\,dx,$$

which is to be maximized subject to the above (isoperimetric) constraints on population and the composite commodity. (Implicit in the above criterion is equal treatment of individuals at the same radius. This is optimal from a more general Benthamite viewpoint if $u(s, z)$ is concave. See Mirrlees [5].) The appropriate Hamiltonian is

$$\mathscr{H} = u\left(\frac{2\pi\theta x}{n(x)}, z(x)\right) n(x) - \psi n(x) - \phi(n(x)\,z(x) + n(x)\,tx + 2\pi\theta x\bar{r}),$$

where ψ and ϕ are costate variables which are constant since the corresponding state variables do not explicitly enter the problem. Hence, assuming an interior solution exists (this can be guaranteed under certain circumstances),

$$\frac{\partial\mathscr{H}}{\partial z} = 0 \qquad \text{so that } u_z = \phi,$$

$$\frac{\partial\mathscr{H}}{\partial n} = 0 \qquad \text{so that } u = u_s s(x) + u_z z(x) + u_z tx + \psi.$$

Substituting these expressions for u and ϕ into the Hamiltonian, it follows that, cancelling a number of terms,

$$\begin{aligned} M &= u_s s(x)\, n(x) - u_z 2\pi\theta x \bar{r} \\ &= 2\pi\theta x u_z \left\{ \frac{u_s}{u_z} - \bar{r} \right\} \\ &= 2\pi\theta x u_z \{ r(x) - \bar{r} \}, \end{aligned}$$

where $r(x)$ is the shadow rent of land in its residential use. Thus the maximized Hamiltonian has the interpretation of Ricardian rent—this is generally true in models of this kind. Now, differentiating the second necessary condition with respect to x, and using the first condition that u_z is constant,

$$\begin{aligned} \frac{du}{dx} &= u_s \frac{ds}{dx} + u_z \frac{dz}{dx} \\ &= \frac{d(u_s)}{dx} s + u_s \frac{ds}{dx} + u_z \frac{dz}{dx} + u_z \cdot t. \end{aligned}$$

Hence

$$\frac{dr(x)}{dx} = \frac{1}{u_z} \cdot \frac{d(u_s)}{dx} = -\frac{t}{s} < 0.$$

The condition corresponding to free choice of R is

$$M|_R = 0 \qquad \text{or} \qquad r(R) = \bar{r}.$$

Hence A.3 of the sufficiency theorem is satisfied and indeed has an intuitively satisfying economic interpretation. (Note that it is not necessarily true here that $dM/dt \leqslant 0$ and hence it is not true a fortiori that $\partial\mathscr{H}/\partial t \leqslant 0$.) Assumption A.2 is trivially satisfied because the state variables do not enter the Hamiltonian. Hence the above necessary conditions are also sufficient for an optimum.

ACKNOWLEDGMENTS

I wish to thank N. V. Long for discussion and an Associate Editor for comments.

REFERENCES

1. K. J. ARROW AND M. KURZ, "Public Investment, the Rate of Return and Optimal Fiscal Policy," John Hopkins Univ. Press, Baltimore, 1970.
2. T. KOOPMANS, Proof for a case where discounting advances the doomsday, *Rev. Econ. Studies*, Symposium (1974), 239–255.
3. E. B. LEE AND L. MARKUS, "Foundations of Optimal Control Theory," Wiley, New York, 1967.
4. O.L. MANGASARIAN, Sufficient conditions for the optimal control of nonlinear systems, *SIAM J. Control* **4** (1966), 139–152.
5. J. MIRRLEES, The optimum town, *Swedish J. Econ.* **74** (1972), 114–135.
6. L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE, AND E. F. MISHCHENKO. "The Mathematical Theory of Optimal Processes," Wiley, New York, 1962.
7. A. SEIERSTAD AND K. SYDSAETER, Sufficient conditions in optimal control theory, *Internat. Econ. Rev.* **18** (1970), 367–392.

JOURNAL OF ECONOMIC THEORY **24**, 446–452 (1981)

A Note on a Closed Input–Output Model with Finite Life-Times and Gestation Lags

MORGAN ÅBERG AND HÅKAN PERSSON

Department of Economics, University of Gothenburg,
Bux 3091, S-400 10 Gothenburg, Sweden

Received September 19, 1978; revised April 7, 1980

INTRODUCTION

The presentation of the dynamic linear input–output model by Lange [4] in 1957 has recently been followed by some studies investigating the properties of such models. See, e.g., Bródy [1] and Johansen [2]. In Lange's model production requirements stem from current use of commodities as well as from growth of stocks.

Bródy [1] in his closed input–output model used disaggregated capital–output coefficients (as did Lange); mathematically, the model is posed as an eigenequation. Recently Johansen [3] has generalized Bródy's model by assuming gestation lags and finite life-time of capital equipment. (The concept of gestation lags in an input–output framework may also be found in an article by Lange [5] written in 1959.)

The purpose of this note is to present an alternative proof and a solution algorithm in a slightly generalized Johansen model. In our model the productive capacity of investment changes by age, but in both models investment may require more than one time period for its construction, and capital goods have different finite life-times.

By a mathematically appropriate formulation of the model, the existence of a unique solution, characterized by balanced growth and full capacity utilization, is easily proved. Furthermore, the mathematical formulation of the non-linear system of equations suggests a simple solution algorithm.

THE MODEL

We consider a closed input–output model where investment is in a sense endogenous. Furthermore, investment is supposed to require more than one

0022-0531/81/030446–07$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

time period for its construction and its productive capacity is supposed to depend on age.

In the model consumption is specified as given input requirements of the household sector and the production net of current use of commodities is by definition all invested. The way of treating investment has as its main characteristics:

(i) Commodity requirements originating from planned capacity expansion is specified as a sequence over the entire gestation time of capacities.

(ii) Capacities have limited and different life-spans and productivity changes due to age.

Let us introduce the following notations:

x_{it}	output of commodity i at time t
k_{it}	gross capacity increase in sector i at time t *(measured by some appropriate index)*
$b_{ij\theta}$	amount of commodity i required at time t per unit gross capacity increase in sector j at time $t+\theta$
a_{ij}	amount of commodity i required at time t per unit of output in sector j at time t
s_j	physical life span of capacity installed in sector j
T_j	maximal gestation time for capacity installed in sector j
$w_{j\tau}$	contribution to net capacity of sector j at time t from a unit of gross capacity installed at $t-\tau$
K_{jt}	total net capacity in sector j at time t
$p_{i,t+\theta}$	price at time t of commodity i occuring at time $t+\theta$

The number of commodities equals the number of sectors and there is no joint production.

Vectors are defined to be positive, semipositive and non-negative, respectively, using the conventions:

$$x > 0 \Rightarrow x_i > 0; \quad x \geqslant 0 \Rightarrow x_i \geqslant 0 \text{ and } x_i > 0 \text{ for some } i; \quad x \geqq 0 \Rightarrow x_i \geqslant 0.$$

In the model considered, output of commodities must equal inputs for each time period. Given the notations above this condition for balance may for each commodity be expressed as

$$x_{it} - \sum_j a_{ij}x_{jt} - \sum_j \sum_{\theta=1}^{T_j} b_{ij\theta}k_{j,t+\theta} = 0, \qquad (1)$$

i.e., total output equals current use of commodities plus commodities stored up to make future capacity increases feasible.

The above formulation makes sense only if the following assumption applies:

$$b_{ij\theta} \geqslant 0 \qquad \text{and some} \qquad b_{ij\theta} > 0. \tag{A1}$$

Over the entire life span, s_j, a unit capacity increase contributes to total net capacity by age determined amounts $w_{j\tau}$. These weights $w_{j\tau}$ can be interpreted to express changes in actual capacity due to aging, e.g., physical depreciation.

Total net capacity in sector j at time t, K_{jt}, is accordingly derived as

$$K_{jt} = \sum_{\tau=0}^{s_j} w_{j\tau} k_{j,t-\tau}. \tag{2}$$

As a second assumption we introduce the following condition

$$w_{j\tau} \geqslant 0 \qquad \text{and some} \qquad w_{j\tau} > 0 \text{ for all } j. \tag{A2}$$

Now, two further conditions are introduced: all capacities are fully utilised and gross capacity increases all grow at the same rate λ.

$$K_{jt} = x_{jt}, \tag{3}$$

$$k_{jt} = (1+\lambda)^{\theta} k_{j,t-\theta}. \tag{4}$$

Relations (2), (3) and (4) can be combined to give

$$x_{jt} = k_{jt} \sum_{\tau=0}^{s_j} w_{j\tau}(1+\lambda)^{-\tau}. \tag{5}$$

From (5) follows that output and capacity increases are proportional by amounts independent of time. Therefore condition (4) implies that output also grows at the uniform and constant rate λ. In growth models a uniform constant rate of growth of output is usually assumed. We give a numerical example in the Appendix to show that such an assumption is not sufficient to guarantee a constant rate of growth of gross capacity in our model. However, the constant rate of growth of gross capacity is needed for the proof given below.

Conditions (1) and (4) give

$$x_{it} - \sum_{j} a_{ij}x_{jt} - \sum_{j} k_{jt} \sum_{\theta=1}^{T_j} b_{ij\theta}(1+\lambda)^{\theta} = 0. \tag{6}$$

Finally, (6) and (5) imply

$$x_{it} - \sum_{j} a_{ij}x_{jt} - \sum_{j} \frac{\sum_{\theta=1}^{T_j} b_{ij\theta}(1+\lambda)^{\theta}}{\sum_{\tau=0}^{s_j} w_{j\tau}(1+\lambda)^{-\tau}} x_{jt} = 0. \tag{7}$$

Dropping time indices and introducing vectors and matrices, (7) is written as

$$x = Ax + B(\lambda)x. \tag{8}$$

A, $B(\lambda)$ are square matrices appropriately defined and x is a column vector.

Now it is possible to formulate a third assumption to be used in this paper:

$$A + B(\lambda) \quad \text{is non-decomposable for} \quad B(\lambda) \geqslant 0. \tag{A3}$$

Economically, this means that every commodity is directly or indirectly necessary for the production of all commodities. A definition of a decomposable matrix can be found in [6, pp. 131].

To make a positive growth rate possible the following final assumption is introduced:

A is productive, i.e., $\rho(A) < 1$, where ρ is the maximum eigenvalue of A. (A4)

As is easily seen, (8) forms an eigenequation for given λ and represents a solution to the generally formulated system

$$\mu x = Ax + B(\lambda)x, \tag{9}$$

where $\mu = 1$. As any other eigenequation (8) has a dual formulation, in this case to be interpreted as a "price formulation"

$$p = pA + pB(\lambda), \tag{10}$$

where p is a row vector of prices.

Using notations in (7) each equation in (10), for every t, may be written

$$\sum_{\tau=0}^{s_j} p_{jt} w_{j\tau} (1+\lambda)^{-\tau} = \sum_{i} \sum_{\tau=0}^{s_i} a_{ij} p_{it} w_{i\tau} (1+\lambda)^{-\tau} + \sum_{i} \sum_{\theta=1}^{l_i} b_{ij\theta} p_{it} (1+\lambda)^{\theta}, \tag{11}$$

which is the investment calculus condition: present value of revenues equals present value of operating costs plus capitalized value of investment goods. Here revenues and costs are all counted per gross capacity unit, and prices are such as to fulfill

$$p_{i,t+\theta} = p_{it}(1+\lambda)^{\theta}, \tag{12}$$

where λ can be interpreted as the rate of interest.

We are interested in solutions to (8) and (10) such that x and p are semipositive, i.e., at least one commodity has to be produced but not necessarily all, and the solution price vector also contains at least one non-zero element. Given that vectors are semipositive, it follows from (4) and (5) that $\lambda > 1$.

As will be proved below there exists a uniquely determined solution with $\lambda > -1$, $x > 0$, ($p > 0$).

Proposition. *With given assumptions* A1–A4 *and* $\lambda > -1$ *there exists a unique solution* λ, x *of* (8) *with* $x > 0$.

Proof. From (7) we see that for $\lambda > -1$ the elements of $B(\lambda)$ are non-negative and continuous functions of λ. From A1 and A2 follows that $\lim_{\lambda \to -1} B(\lambda) = 0$. For all j, $w_{j\tau} > 0$ for some τ. If for some i and j, $b_{ij\theta} = 0$ for all $\theta = 1, \ldots, T_j$, then the corresponding element $b_{ij}(\lambda)$ of the matrix $B(\lambda)$ is equal to zero for all λ. For the non-zero elements of $B(\lambda)$ we have $b_{ij}(\lambda) > 0$ for $\lambda > -1$, $\lim_{\lambda \to \infty} b_{ij}(\lambda) = \infty$ and

$$\frac{db_{ij}}{d\lambda} = \frac{\left(\sum_{\tau=0}^{S_j} w_{j\tau}(1+\lambda)^{-\tau}\right)\left(\sum_{\theta=1}^{T_j} b_{ij\theta}\theta(1+\lambda)^{\theta-1}\right)}{\left(\sum_{\tau=0}^{S_j} w_{j\tau}(1+\lambda)^{-\tau}\right)^2} - \frac{\left(\sum_{\theta=1}^{T_j} b_{ij\theta}(1+\lambda)^{\theta}\right)\left(\sum_{\tau=0}^{S_j} s_j(-\tau)(1+\lambda)^{-\tau-1}\right)}{\left(\sum_{\tau=0}^{S_j} s_j(1+\lambda)^{-\tau}\right)^2}$$

is positive for $\lambda > -1$. Thus, these elements are strictly increasing functions of λ. Since there are i and j such that $a_{ij} + b_{ij}(\lambda) \to \infty$ as $\lambda \to \infty$ and since $A + B(\lambda)$ is indecomposable $\lim_{\lambda \to \infty} \mu = \infty$ in (9).

Due to the productivity of A there is $\lambda > -1$ such that $\mu(\lambda) < 1$. Moreover, $\lambda' > \lambda$ implies that $\mu(\lambda') > \mu(\lambda)$ (see [6, p. 135]). μ is a continuous function of the elements of $B(\lambda)$ (see [7]) and hence μ is a continuous function of λ. Thus, in the interval $(-1, \infty)$ there is a unique λ satisfying (8). Due to the indecomposability assumption the corresponding vector $x > 0$ is unique up to the multiplication by scalars. Q.E.D.

It easily follows that there also exists a unique solution λ, p of (10) with $p > 0$, in the sense of the proposition, and that the rates of growth and interest coincide.

Note that (A4) is necessary but not sufficient to guarantee a positive growth rate. The model requires a net production of investment commodities as long as $\lambda > -1$. It easily follows from the proposition that a necessary and sufficient condition for a positive growth rate is that the matrix $(A + B(0))$ is productive. The finite life time of capacities requires a production for investment to maintain net capacities. Hence, if the economy is not so productive as to secure a zero growth rate, a positive growth rate will also be impossible.

SOLUTION ALGORITHMS

From the proof of the proposition follows that in

$$\mu x = (A + B(\lambda))x.$$

μ is a continuous strictly increasing function of λ.

For a given λ we calculate the maximal eigenvalue μ and the corresponding eigenvector x with the aid of the power method (see [8]).

When μ has been calculated we make an adjustment of λ based on the deviation of μ from 1 as

$$\lambda^{n+1} = \lambda^n + K(1 - \mu^n),$$

where K is a positive constant.

With this new value of λ we again use the power method in finding μ and the process continues. It can be formally proved that the algorithm works, i.e., $\mu(\lambda^n) \to 1$ and for the limit $x = Ax + B(\lambda)x$.

APPENDIX

We will here give an example that shows that with given assumptions a constant rate of growth of output is not sufficient to guarantee that increases of gross capacity grow at a constant rate.

We use a two-sector model with an input–output matrix

$$A = \begin{pmatrix} 0.5 & 0.2 \\ 0.4 & 0.6 \end{pmatrix}.$$

B^1 and B^2 denote the matrices with elements $b_{ij\theta}$ for $\theta = 1$ and $\theta = 2$, respectively.

$$B^1 = B^2 = \begin{pmatrix} 0.2 & 0.1 \\ 0 & 0 \end{pmatrix}.$$

The life span of capacity is two periods for both sectors, i.e., $s_1 = s_2 = 1$ in (2) and $w_{j\tau} = 1, j = 1, 2, \tau = 0, 1$.

A solution of the model is $x_{1t} = x_{2t} = 1$ for all t, i.e., $\lambda = 0$. Gross capacity increases are such that $k_{jt} + k_{j,t+1} = 1, j = 1, 2$ and all t. We can then for instance have a sequence of gross capacity increases such that

$$\begin{aligned} k_{jt} &= 3/4 \qquad t \text{ odd}, \\ &= 1/4 \qquad t \text{ even}, \end{aligned}$$

consistent with a balanced growth of output.

ACKNOWLEDGMENTS

We are grateful to Leif Johansen for helpful talks in the preparation of this note.

REFERENCES

1. A. BRÓDY, "Proportions, Prices and Planning," North–Holland, Amsterdam, 1970.
2. L. JOHANSEN, The rate of growth in dynamic input–output models: Some observations along lines suggested by O. Lange and A. Bródy, *in* "Jahrbuch der Wirtschaft Osteuropas," Band 4, 68–69, Günter Olzog Verlag, Munich, 1973.
3. L. JOHANSEN, On the theory of dynamic input–output models with different time profiles of capital construction and finite life-time of capital equipment, *J. Econ. Theory* **16** (1978), 513–533.
4. O. LANGE, Some observations on input–output analysis, *in* "Papers in Economics and Sociology," Pergamon, New York and PWN. Warsaw, 1970. Originally published in 1957.
5. O. LANGE, Produkcyjno-techniczne podstawy efektywności inwestycji, *in* "Pisma ekonomiczne i spoleczne," PWN, Warsaw, 1961. Originally published in 1959.
6. H. NIKAIDO, "Introduction to Sets and Mappings in Modern Economics," North–Holland, Amsterdam, 1970.
7. A.M. OSTROWSKI, "Solution of Equations in Euclidian and Banach Spaces," 3rd ed., Academic Press, 1973.
8. J. H. WILKINSON, "The Algebraic Eigenvalue Problem," Oxford Univ. Press (Clarendon), London, 1965.

Bounds on the Solutions to $x' = SAx$ where A Is S-Stable

D. J. HARTFIEL

Mathematics Department,
Texas A & M University, College Station, Texas 77843

Received January 31, 1980; revised April 7, 1980

Stable matrices by definition are square matrices whose eigenvalues have negative real parts. Such matrices have been studied extensively since the turn of the century. More recently, studies on markets for finitely many commodities have lead economists to consider more general notions of stability. A first such notion was given by Enthoven and Arrow in [3]. A modern version of their notion is that a real square matrix A is D-stable if and only if DA is stable for all positive definite diagonal matrices D. This notion was generalized by Arrow and McManus in [1] by defining a real square matrix A to be S-stable if and only if SA is stable for all positive definite real symmetric matrices S.

Interest in these, and other such notions of stability, arise by considering the differential equation

$$x' = SAx,$$

$$x(0) = a,$$

where S is a positive definite real symmetric matrix. This equation describes the price behavior in a multiple commodity market. As stated by Metzler in [6, p. 284], for the case when S is a diagonal matrix, economists are usually more confident in their knowledge of the entries of A than they are in their knowledge of the entries of S. Thus, it is desirable to describe the behavior of the solution to the differential equation above in terms which are independent of S. Thus, if A is S-stable, $\text{limit}_{t\to\infty} x(t) = 0$ independent of S.

Although S-stability is useful in describing the long run behavior of x, independent of S, it says little about the solution over finite intervals. To demonstrate this, we give the following example.

EXAMPLE. Consider

$$x' = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} x,$$

$$x(0) = a, \quad \text{with} \quad |a| = 1.$$

0022-0531/81/030453-05$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved.

By using a characterization of S-stability developed by David Carlson in [2, p. 62] it can be shown that $A = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ is S-stable. Thus, $\text{limit}_{t\to\infty} x(t) = 0$ for any fixed $d_1 > 0$ and $d_2 > 0$. However, it can also be shown here that for any $t > 0$, $d_1 > 0$ and $d_2 > 0$ can be chosen so that $|x(t)|$ is arbitrarily close to $(t^2 + 1)^{1/2}$. Thus, even though $\text{limit}_{t\to\infty} x(t) = 0$ for any fixed $d_1 > 0$ and $d_2 > 0$, it is also true that, for some $d_1 > 0$ and $d_2 > 0$, $|x(t)|$ can be quite large for large values of t.

Thinking about the above example leads one to consider finding an exponential bound, in terms of S and A, on $|x(t)|$. Such a bound could be of use in studying the rate of convergence of $x(t)$.

We begin with a brief discussion of the set S_n of $n \times n$ S-stable matrices. These matrices were characterized by Carlson in [2, p. 62] as follows:

LEMMA. *A real square matrix A is S-stable if and only if*

(1) *$x'Ax \leqslant 0$ for all real vectors x,*

(2) *A is nonsingular, and*

(3) *for each pair of real vectors x and y for which $(A + A')x = 0$ and $(A + A')y = 0$ it follows that $x'Ay = 0$.*

As a consequence of this lemma, we have a characterization of the interior of S_n.

COROLLARY. *Let $A \in S_n$. Then A is in the interior of S_n if and only if $x'Ax < 0$ for all real vectors $x \neq 0$.*

Square matrices A for which $x'Ax < 0$, for all $x \neq 0$, are called negative quasi-definite. An exponential bound on $x(t)$ for this class of matrices follows.

THEOREM. *Let S be an $n \times n$ positive definite real symmetric matrix and A a negative quasi-definite matrix. Then the solution $x(t)$ to*

$$x' = SAx,$$
$$x(0) = a,$$

satisfies $|x(t)| \leqslant (\lambda_1(S)/\lambda_n(S))^{1/2} |a| e^{st}$, where $\lambda_1(S) \leqslant \cdots \leqslant \lambda_n(S)$ are the eigenvalues of S and $s = \max_{x\neq 0} x'Ax/x'S^{-1}x < 0$.

Proof. Let x be the solution to

$$x' = SAx,$$
$$x(0) = a.$$

Then

$$S^{-1}x' = Ax$$

and

$$x^tS^{-1}x' = x^tAx$$

or

$$(x^tS^{-1}x)' = 2x^tAx.$$

Integrating both sides yields

$$x^tS^{-1}x = a^tS^{-1}a + \int_0^t 2x^tAx\,d\tau$$

so that

$$x^tS^{-1}x \leqslant a^tS^{-1}a + 2s\int_0^t x^tS^{-1}x\,d\tau.$$

Thus by Gronwall's inequality [5, p. 24],

$$x^tS^{-1}x \leqslant (a^tS^{-1}a)\,e^{2st}.$$

Since $x^tS^{-1}x \geqslant \lambda_1(S^{-1})x^tx$ and $a^tS^{-1}a \leqslant \lambda_n(S^{-1})a^ta$, by Rayleigh's principle [4, p. 144], it follows that

$$x^tx \leqslant \frac{\lambda_n(S^{-1})\,a^tae^{2st}}{\lambda_1(S^{-1})}$$

or

$$|x(t)|^2 \leqslant \frac{\lambda_n(S)}{\lambda_1(S)}|a|^2\,e^{2st}. \quad \blacksquare$$

A bound on the number s may be computed as follows.

COROLLARY 1. *Let S be an $n \times n$ positive definite real symmetric matrix and A a negative quasi-definite matrix. Then*

$$s = \max_{x \neq 0}\frac{x^tAx}{x^tS^{-1}x} \leqslant \frac{1}{2}\lambda_1(S)\,\lambda_n(A + A^t).$$

Thus if

$$x' = SAx,$$
$$x(0) = a,$$

then

$$|x(t)| \leqslant \left(\frac{\lambda_1(S)}{\lambda_n(S)}\right)^{1/2} |a|\, e^{(1/2)\lambda_1(S)\lambda_n(A+A')t}.$$

Proof. The proof follows by noting that

$$\begin{aligned} s &= \max_{x \neq 0} \frac{x'Ax}{x'S^{-1}x} \\ &\leqslant \frac{\max_{|x|=1} x'Ax}{\max_{|x|=1} x'S^{-1}x} \\ &\leqslant \frac{\frac{1}{2}\max_{|x|=1} x'(A+A')x}{\max_{|x|=1} x'S^{-1}x} \end{aligned}$$

andby Rayleigh's principle

$$\begin{aligned} &\leqslant \frac{\frac{1}{2}\lambda_n(A+A')}{\lambda_n(S^{-1})} \\ &\leqslant \tfrac{1}{2}\lambda_1(S)\,\lambda_n(A+A'). \quad \blacksquare \end{aligned}$$

The theorem may be applied to D-stable matrices as follows.

COROLLARY 2. *Let D be an $n \times n$ positive definite diagonal matrix and A a negative quasi-definite matrix. Then the solution $x(t)$ to*

$$\begin{aligned} x' &= DAx, \\ s(0) &= a, \end{aligned}$$

satisfies $|x(t)| \leqslant (\max_{i,j} d_i/d_j)^{1/2}\, |a|\, e^{st}$.

Finally, for any n, if A is an $n \times n$ real negative definite symmetric matrix, and $S = I$ then $s = \lambda_n(A)$ and the bound of the theorem yields

$$|x(t)| \leqslant |a|\, e^{\lambda_n(A)t}.$$

Since, in this case,

$$x(t) = \frac{(a'u^1)}{a'a} e^{\lambda_1(A)t} u^1 + \cdots + \frac{(a'u^n)}{a'a} e^{\lambda_n(A)t} u^n,$$

where $u^1,\ldots,u^n$ is an orthonormal set, it follows, by taking $a=u^n$, that

$$x(t)=e^{\lambda_n(A)t}u^n$$

so

$$|x(t)|=e^{\lambda_n(A)t}.$$

Hence the bound is achieved at any real negative definite symmetric matrix A for some poisitive definite symmetric matrix S.

REFERENCES

1. K. ARROW AND M. MCMANUS, A note on dynamic stability, *Econometrica* **26** (1958), 448–454.
2. D. CARLSON, A new criterion of H-stability of complex matrices, *Linear Algebra Appl.* **1** (1968), 59–64.
3. A. C. ENTHOVEN AND K.J. ARROW, A theorem on expectations and the stability of equilibrium, *Econometrica* **24** (1956), 288–293.
4. J. N. FRANKLIN, "Matrix Theory," Prentice–Hall, Englewood Cliffs, N. J., 1968.
5. P. HARTMAN, "Ordinary Differential Equations," Wiley, New York, 1964.
6. L. A. METZLER, Stability of multiple markets: The Hick's condition, *Econometrica* **13** (1945), 277–292.

JOURNAL OF ECONOMIC THEORY **24**, 458 (1981)

Announcement

A New Editorial Policy

Effective May 1, 1981, *all* manuscripts forwarded to the publisher from the *Journal of Economic Theory* editorial office must include a brief abstract to be published at the beginning of the article. Published abstracts are useful to the various scholarly indexing services.

Authors are advised to consult our recently revised "Information for Authors" on the inside back cover. For further information, write to Susan W. Schulze, Assistant Editor, *Journal of Economic Theory*, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia PA 19104 USA.

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved.

Papers to Appear in Forthcoming Issues

The following articles are among those newly accepted for publication in the *Journal of Economic Theory*. They are listed alphabetically by first author.

Jordi Andreu, Rationalization of Market Demand on Finite Domains. Correspondent: Professor Jordi Andreu, Department of Economics, Universidad Autonoma de Barcelona, Ballaterra (Barcelona), Spain.

Yves Balasko, Equilibria and Efficiency in the Fixprice Setting. Correspondent: Professor Yves Balasko, CEPREMAP, 142, rue du Chevaleret, 75013 Paris, France.

Mario Girardi and Giorgio Israel, Some Quasi-Globally Stable Processes of Price Adjustment. Correspondent: Professor Mario Girardi, Istituto Matematico "G. Castelnuovo," Università degli Studi di Roma, Città Universitaria, 00100 - Roma, Italy.

Gur Huberman, A Simple Approach to Arbitrage Pricing Theory. Correspondent: Professor Gur Huberman, Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, IL 60637.

Martine Quinzii, An Existence Theorem for the Core of a Productive Economy with Increasing Returns. Correspondent: Professor Martine Quinzii, Laboratoire d'Econometrie, L'Ecole Polytechnique, 17 rue Descartes, 75005 Paris, France.

Jan Svejnar, On the Theory of a Participatory Firm. Correspondent: Professor Jan Svejnar, Department of Economics, Cornell University, 486 Uris Hall, Ithaca, N.Y. 14853.

Ferenc Szidarovsky and Sidney J. Yakowitz, Contributions to Cournot Oligopoly Theory. Correspondent: Professor Sidney Yakowitz, Systems and Industrial Engineering, University of Arizona-Tucson, Tucson, AZ 85721.

Makoto Yano, A Note on the Existence of an Optimal Capital Accumulation in the Continuous Time Horizon. Correspondent: Professor Makoto Yano, Department of Economics, Cornell University, Ithaca, NY 14853.

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

JOURNAL OF ECONOMIC THEORY **24**, 460 (1981)

Author Index for Volume 24

A

ÅBERG, MORGAN, 446

B

BALASKO, YVES, 112, 143
BARON, DAVID P., 70
BEWLEY, TRUMAN F., 265
BURDETT, KENNETH, 362
COUGHLIN, PETER, 226

D

DE BONDT, RAYMOND R., 70

F

FIELD, C. A., 168
FOURGEAUD, CLAUDE, 413
FREIXAS, XAVIER, 296

H

HARTFIEL, D. J., 453

K

KANEKO, MAMORU, 311
KEIDING, HANS, 393
KHAN, M. ALI, 218
KIM, KI HANG, 40

L

LAITNER, JOHN, 377
LENCLUD, BERNARD, 413
LEWIS, GEOFFREY, 240
LIM, CHIN, 328

M

MACLEAN, L. C., 168
MALUEG, DAVID A., 362
MICHEL, PHILIPPE, 413
MILNE, FRANK, 205
MOULIN, HERVÉ, 398

N

NISHIMURA, KIYOHIKO GIICHI, 157
NITZAN, SHMUEL, 226
NTI, KOFI O., 187

P

PERSSON, HÅKAN, 446

R

ROBSON, ARTHUR J., 437
ROUSH, FRED W., 40

S

SATO, FUMITAKA, 48
SHELL, KARL, 112, 143
SHUBIK, MARTIN, 187
SUTHERLAND, W. R. S., 168

W

WILSON, CHARLES A., 95

Y

YAARI, MENAHEM E., 1
YAMAZAKI, AKIRA, 218

Printed in Belgium

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

JOURNAL OF ECONOMIC THEORY

Volume 25, Number 1, August 1981

Copyright © 1981 by Academic Press, Inc.
All Rights Reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc. (21 Congress Street, Salem, Massachusetts 01970), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre-1981 articles are the same as those shown for current articles.

Published bimonthly at 41 Tempelhof, Bruges, Belgium by
Academic Press, Inc., 111 Fifth Avenue, New York, N.Y. 10003.

1981: Volumes 24–25. Price $138.00 U.S.A.; $165.00 outside U.S.A.
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publisher. All correspondence and subscription orders should be addressed to the office of the Publishers at 111 Fifth Avenue, New York, N.Y. 10003. Send notices of change of address to the office of the Publishers at least 6 to 8 weeks in advance. Please include both old and new addresses. Postmaster, send changes of address to *Journal of Economic Theory*, 111 Fifth Avenue, New York, New York 10003.

Second class postage paid at Jamaica, N.Y. Air freight and mailing in the U.S.A. by Publications Expediting, Inc., 200 Meacham Avenue, Elmont, New York 11003.

Copyright © 1981 by Academic Press, Inc.
Printed in Bruges, Belgium, by the St. Catherine Press, Ltd.

Optimal Adaptive Price Search

DONALD B. ROSENFIELD

Arthur D. Little, Cambridge, Massachusetts

AND

ROY D. SHAPIRO

Graduate School of Business Administration,
Harvard University, Boston, Massachusetts 02163

Received March 30, 1978; revised December 3, 1979

INTRODUCTION

Consider the problem of a consumer desiring to buy a good which is sold at a variety of prices. Such a buyer will be able to obtain a sequence of price quotations from various sellers. When should the buyer stop searching for a lower price? If he continues, he may obtain a lower price, but he incurs additional time, effort and expense in searching. The optimal price search decision is one of balancing the expectation of a possibly lower price obtained by continued search against cumulative costs from search. When the prices arise from a known probability distribution, the optimal search strategy has certain well-defined and desirable properties. In the more realistic case that the buyer has only partial information about the price distribution and must update his beliefs after observing prices in the marketplace, the problem becomes more complex, and search strategies become less well behaved.

The major purpose of this article is to determine new conditions under which optimal policies have these desirable characteristics in the case where the consumer possesses only incomplete information about the distribution of prices. Price quotations he receives, in addition to providing him with price possibilities, provide him with information about that distribution. Rothschild [15], in citing an unpublished paper of Gastwirth, has noted that small specification errors for an assumed distribution can lead to substantial changes in expected number of searches and expected cost. Consequently, the assumption of a known distribution of prices can have a profound effect on total cost.

0022-0531/81/040001-20$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

Our approach is Bayesian in nature. Our intent is to determine when certain properties (such as the existence of optimal reservation prices) hold when there is Bayesian updating of the consumer's prior distribution of prices. DeGroot [5], Rothschild [15] and Albright [2] presented some results along these lines. Telser [17] employed simulation to examine the properties and costs of certain reasonable search rules when the distribution is unknown. His analysis examines the expected gain from one more search; this may not be optimal. In addition, since a future composite distribution of prices depends on the next price drawn, the question of stopping on the next price becomes $p \leqslant a(p)$. The region where this inequality holds may be disjoint in p. Rothschild recognized this problem and discussed the problem of price reservations independent of p; Kohn and Shavell [7] were also cognizant of this issue. They note that their "switchpoints" depend on an updated distribution after observation.

Our results generalize regularity conditions, which hold if the distribution is known, to the Bayesian case. For the situation in which the buyer is not allowed to accept prices previously quoted (*no recall allowed*), we derive conditions under which the optimal policy is *reservation*. That is, the buyer accepts a price if and only if it is below the reservation. For the *recall allowed* case, we are interested in the existence and implications of myopic policies. (Chow and Robbins [4], Abdel-Hameed [1], and Pratt *et al.* [13] have also examined this question.) Before stating our results, we shall briefly review the characteristics of optimal price search under a known distribution.

II. PROPERTIES OF OPTIMAL STOPPING RULES UNDER A (KNOWN) FIXED DISTRIBUTION OF PRICES

Recall Allowed

Assume prices are distributed according to a known distribution F with finite mean and variance.[1] At any stage, the state of the search may be represented by p^*, the minimum price observed thus far. Note that, under the assumption of a known distribution of prices, p^* contains all information relevant to future decisions.

Under *any* known distribution of prices, the optimal stopping rule in both finite- and infinite horizon case has three important desirable and well-known properties. First, it is *myopic*. That is, the optimal policy can be determined by a one-period look ahead, so that after a price quotation, the consumer can view the problem as one in which he is allowed only one more price quote regardless of the actual number of quotes which might be available in the future. Secondly, it is characterized by a *reservation price*, i.e., a price P, independent of the number of periods remaining or originally available, such that the buyer should continue (if possible) until he obtains a price quotation less than or equal to P. Thirdly, although recall is allowed, the buyer will never return to a previous price quotation unless he reaches the horizon. Also, the expected infinite horizon cost exists and equals the limit of the finite horizon costs. These results follow directly from the dynamic programming formulation for the minimum expected cost with one period to go. The gain from observing an additional price in the remaining time period is $G(p^*) - c$, where c is the cost of obtaining a price quotation and $G(p^*) = \int_{-\infty}^{p^*} F(p)\,dp$. Since $G(p^*)$ is monotonically increasing in p^*, once the expected gain from one more search becomes negative, it stays negative. If the gain is positive, the buyer should certainly continue at least for one more period. Thus for any horizon, the buyer should continue if and only if $G(p^*) \geqslant c$ or $p \leqslant G^{-1}(c)$ since G is monotonic.

---

[1] We actually need require only that $\min(0, p)$ has finite variance. If $c > 0$ and the price distribution has a finite mean and variance, then the existence of an optimal stopping policy for both this case and the case of no recall allowed has been established by DeGroot [5] and Yahav [18].

No Recall Allowed

With no recall allowed, the future return if the buyer continues cannot depend on the current price. Let the current price p be the state, sufficient because under a known price distribution, trials are independent, and the only information necessary to decide whether to stop or to continue is the terminal value received when stopping. Observe that the minimum expected t-period cost is

$$V_t(p) = \min \left\{ p, c + \int_{-\infty}^{\infty} V_{t-1}(p)\,dF(p) \right\};$$

an analogous expression holding for the infinite horizon case. It follows then, that the optimal policy is reservation but not necessarily myopic. The reservations for finite horizon problems converge to the infinite horizon reservation. Moreover, the infinite horizon reservation price and optimal return have values identical to those obtained for the recall allowed case, since min $(p, G^{-1}(c))$ satisfies both sets of infinite horizon recursive equations and characterizes a reservation price $G^{-1}(c)$.

III. Unknown Price Distribution — Examples

If we do not assume that the buyer is aware of the distribution of prices, and instead updates his beliefs about that distribution after observing prices from the marketplace, none of the desirable properties described in Section II necessarily hold. Consider the following examples.

Note that this condition tests the relative value of stopping for the one-period look-ahead. If (2) determines the optimal policy for the infinite horizon or the t-period problem, then the optimal policy is myopic. Observe that (2) is equivalent to

$$c - G(p_N^* | \mathbf{p}) > 0, \tag{3}$$

where

$$G(x | \mathbf{p}) = \int_{\infty}^{x} F(\rho | \mathbf{p})\, d\rho,^{3}$$

which can be viewed as the gain from one-period look-ahead. An optimal stopping rule determined by (3) is thus myopic. A condition sufficient to guarantee a myopic optimal stopping rule follows.

THEOREM 1. *If $G(p_N^* | \mathbf{p})$ never crosses c from below with additional observations, then the optimal policy is myopic.*

Remarks. (1) The results hold for c varying with N and a terminal value which is a function of p_N^*(and N). In this case, G still represents the one-period gain.

(2) Under the hypothesis of Theorem 1, if $G(p_N^* | \mathbf{p}) < c$, it stays less than c. That is,

$$G(p_N^* | \mathbf{p}) < c \Rightarrow G(p_N^* \wedge p_{N+1} | \mathbf{p}, p_{N+1}) < c$$

for possible observations p_{N+1} and all previous sequences of prices.

Proof (by induction). For one period the assertion is self-evident. The induction hypothesis states that the assertion holds for the $(t-1)$-period problem. If myopia says stop, it will say so again if p_{N+1} is observed. That is, if (3) holds for $\mathbf{p}$ then, by the hypothesis of the theorem, (3) will hold for $(\mathbf{p}, p_{N+1})$, and the optimal action with $t-1$ periods to go is stop. Thus,

$$V_{t-1}(\mathbf{p}, p_{N+1}) = p_{N+1} \wedge \mathbf{p}_N^*.$$

In this case, for the t-period problem,

$$c + \int V_{t-1}(\mathbf{p}, \rho)\, dF(\rho | \mathbf{p}) - p_N^* = c + \int (\rho \wedge p_N^*)\, dF(\rho | \mathbf{p}) - p_N^* > 0.$$

---

[3] $G(p_N^* | \mathbf{p})$ is a function which depends on the first N observations and is evaluated at p_N^*.

Thus, by the equivalence of (2) and (3), stopping is optimal. If (3) does not hold,

$$c + \int V_{t-1}(\mathbf{p}, \rho)\, dF(\rho|\mathbf{p}) - p_N^*$$

$$\leqslant c + \int (\rho \wedge p_N^*)\, dF(\rho|\mathbf{p}) - p_N^* \leqslant 0$$

and continuing is optimal. Thus the desirability of continuing can be determined by (3) (hence, equivalently, by (2) which is the myopic one-period look-ahead condition).

If G (the expected gain from observing one more price) never crosses c from below, then as soon as the expected gain from one more observation becomes less than the cost incurred in obtaining it, we are assured that any future gain from observation will also be less than its costs. In this case, as *we have demonstrated above, a myopic decision rule is appropriate.* Without constraining G to at most one crossing, we cannot assert that a myopic policy is optimal: it might happen that the expected gain from one more observation would not cover its cost, but that the expected gain from a second observation would be greater than the two costs combined. In the absence of conditions on the behavior of G, analysis that examines the expected gain of one more search is inadequate for the determination of an optimal policy.

We use the above condition on G to give three examples of prior price distributions which have myopic policies. For the first two, we show that G is monotonically decreasing in additional observations, thereby guaranteeing a myopic optimal policy for any search cost.

Example 1: Multinomial Distribution with a Dirichlet Prior

Intuitively, such a prior updates a multinomial such that the probability of the price just observed increases a little and the probabilities of other prices decrease slightly. Formally, one can characterize the distribution at any time by parameters $N_1 \cdots N_m$, which are analogous to frequencies of prices $p_1, \ldots, p_m$. The probability of observing price p_i is N_i/N, where $N = \sum N_i$. After any observation of price p_i the new N_i is updated to $N_i + 1$. This distribution has characteristics which seem consistent with consumer behavior.

$$\begin{aligned} G(p_{N+1}^* | p_1, \ldots, p_{N+1}) &= (N/N+1)\, G(p_{N+1}^* | p_1, \ldots, p_N) \\ &\leqslant (N/N+1)\, G(p_N^* | p_1, \ldots, p_N). \end{aligned} \tag{4}$$

As an example, consider three possible prices: 1, 2, and 3 with prior

$$P(1) = \tfrac{1}{6}, \qquad P(2) = \tfrac{1}{3}, \qquad P(3) = \tfrac{1}{2},$$
$$c = 0.55, \qquad N = 6.$$

If the first observation yields 3, we continue, but *whatever* we observe in the next period, we will stop since $G(i\,|\,3, i) < c$. This may seem counter-intuitive, but it simply says that after the first 3 is drawn, we are not yet quite sure if the distribution is "bad" enough to give up and accept 3; if we draw another 3, we will be. In general, if

$$\frac{N}{N+1} G(p_N^* \,|\, \mathbf{p}) < c < G(p_N^* \,|\, \mathbf{p}),$$

then the reservation price is $\leqslant p_N^*$, and the optimal policy at stage $N+1$ is stop, regardless of the value of p_{N+1}.

It follows from (4) that the multinomial/Dirichlet satisfies the following corollary:

COROLLARY 1. *If* $G(p_N^* \,|\, p_1, \ldots, p_N)$ *depends only on* N *and* p_N^*, *and is non-decreasing in* p_N^* *for* $N \geqslant 1$ *and non-increasing in* N, *then the optimal policy is reservation in terms of the best price so far. Furthermore, the reservations are non-decreasing in* N.

Remarks. The monotonicity of reservations in N can cause the situation illustrated above, where the reservation becomes larger than the best price and stopping is optimal under any circumstances. Note also that this corollary covers the case of F known, since then G depends only on p_N^*.

Proof. Since $p_0^* = \infty$, one must choose a first observation for $N \geqslant 1$; from the monotonicity of G in p_N^* and N, G crosses c at most once. The optimal policy is hence myopic and the optimal action is determinable from the sign of

$$G(p_N^* \,|\, \mathbf{p}) - c.$$

Since G is monotonic in p_N^* one should stop if

$$p_N^* \leqslant \inf\{p\colon G(p \,|\, \mathbf{p}) - c \geqslant 0\} = P_N$$

and go on otherwise. As G is monotonic in N, the reservation price P_N is non-decreasing in N.

Our second example also satisfies the hypothesis of the corrollary.

*Example 2: **Exponential with Unknown Origin***[4]

In this example the distribution of prices is $f(p)=e^{-a(p-t)}$, where t is the unknown origin of the distribution. This distribution is realistic in the sense that the buyer does not often know what the lower bound on prices is. If, however, the buyer feels that prices are clustered near the lower bound, this exponential model is appropriate. Any prior density $r(t)$ with $\log r(t)$ concave may be used. (This includes the exponential and, as special case, the uniform.) It then follows that the hypothesis of Corollary 1 is satisfied and that the optimal policy is myopic with increasing reservations. Verification of this and a generalization is given in the Appendix.

Example 3: Normal with Unknown Mean

This final example does not satisfy the corollary; it is due to DeGroot [5]. Assume that the distribution of prices is Normal with unknown mean μ and variance 1, and that the prior on μ is Normal with mean μ_0 and variance $1/\tau$ (precision τ). Then if

$$[(\tau+1)/(\tau+2)]\, c^2 \geqslant 1/2\pi$$

the optimal policy with recall allowed is myopic. Furthermore, the optimal stopping rule is to stop when

$$p_N^* - E(p\,|\,p_1,\dots,p_N) \leqslant \frac{\psi^{-1}\{c[(\tau+N)/(\tau+N+1)]^{1/2}\}}{[(\tau+N)/(\tau+N+1)]^{1/2}},$$

where

$$\psi(x)=\int_{-\infty}^{x} \Phi(\xi)\,d\xi,$$

Φ being the standardized Normal cumulative distribution function.

The proof is similar to DeGroot's and is omitted.

Notice that this policy is not necessarily reservation. The region of termination as a function of p_N would include very small p_N and very large p_N(where $E(p|p_1,\dots,p_N)$ is very large—in this case the buyer suddenly realizes prices are much higher than he believed.)

Conceptually, since G is a weighted combination of the mass on "low" prices, Theorem 1 says that for an optimal policy to be myopic, the price distribution must be such that observing "high" prices (i.e., prices greater than p^*) causes the mass associated with "low" prices (i.e., pices below p^*) to decrease. For example, it can be shown a Normal price distribution with unknown variance is characterized by a non-myopic optimal search strategy, *since observing high prices increases the likelihood of subsequently observing low prices. In general, we would expect not to find myopic optimal policies* associated with distributions whose unknown parameter is spread related. *Conversely, if the price distribution is such that a new price observation causes a translation in the distribution without increasing the mass below the* observation, then we might expect to find the distribution characterized by a myopic optimal search strategy. Distributions characterized by origin-related (or, in some cases, mean-related) unknown parameters such as *those* discussed in Examples 2 and *3 typically have this property.*

---

[4] This example was suggested by John Pratt.

V. UNKNOWN PRICE DISTRIBUTION—NO RECALL ALLOWED

The important issue in this case is whether the optimal policy is a reservation price policy. Rothschild [15] examined this problem and showed that the optimal policy is reservation for a multinomial distribution with the Dirichlet prior. Our result, although not covering the Dirichlet-multinomial specifically, is somewhat more general.

We look at two buyers who receive price quotations in the marketplace. Whenever their experiences (i.e., their vectors of observed prices) differ in only one quote, we would like their respective distributions for the next price to be "not too different."

DEFINITION. Given a vector of observed prices, let a *continuation price* be any price which, if observed next, would lead the buyer, acting under the optimal policy, to continue searching. Similarly, observing a *stopping price* would cause the buyer to stop. If the indicated action is strictly preferred over the alternative, we call the price a *strict* stopping or continuation price.

THEOREM 2. *Suppose that whenever* $\mathbf{p} \geqslant \mathbf{q}$ *are two N-vectors of observed prices which differ in at most one component,*

$$F(x \mid \mathbf{p}) \leqslant F(x \mid \mathbf{q}) \qquad \text{for all } x$$

and

$$G(x \mid \mathbf{q}) \leqslant G\left(x + \frac{\Delta}{N} \mid \mathbf{p}\right) \qquad \text{for all } x,$$

where Δ *denotes the positive component difference between* $\mathbf{p}$ *and* $\mathbf{q}$. *Then*

1. *The difference in the expectations of the next price observed is bounded by* 0 *and* Δ/N, *i.e.,*

$$0 \leqslant E_{\rho\mid\mathbf{p}}(\rho) - E_{\rho\mid\mathbf{q}}(\rho) \leqslant \Delta/N. \tag{1}$$

2. *The difference in the expectations of continued search is bounded by 0 and Δ/N, i.e.,*

$$0 \leqslant E_{\rho|\mathbf{p}} V_t(\mathbf{p}, \rho) - E_{\rho|\mathbf{q}} V_t(\mathbf{q}, \rho) \leqslant \Delta/N \tag{2}$$

for all t (including the infinite horizon). In particular, then, if p_N and q_N are continuation prices given $p_1, \ldots, p_{N-1}$ and $q_1, \ldots, q_{N-1}$, respectively, then

$$0 \leqslant V_t(\mathbf{p}) - V_t(\mathbf{q}) \leqslant \Delta/N.$$

3. *The optimal policy is characterized by a sequence of reservation prices.*

Proof.

$$\begin{aligned} E_{\rho|\mathbf{p}}(\rho) - E_{\rho|\mathbf{q}}(\rho) - \Delta/N &= \int_{-\infty}^{\infty} F(\rho \mid \mathbf{q})\, d\rho - \int_{-\infty}^{\infty} F(\rho \mid \mathbf{p})\, d\rho - \Delta/N \\ &\leqslant \lim_{T \to \infty} \int_{-\infty}^{T} F(\rho \mid \mathbf{q})\, d\rho - \int_{-\infty}^{T+\Delta/N} F(\rho \mid \mathbf{p})\, d\rho \\ &= \lim_{T \to \infty} G(T \mid \mathbf{q}) - G(T + \Delta/N \mid \mathbf{p}) \leqslant 0, \end{aligned}$$

thus establishing (1). (Non-negativity follows directly from stochastic dominance.) We now show conclusions 2 and 3 of the theorem by induction on the number of periods remaining. We use the recursive expression

$$V_t(\mathbf{p}) = \min\left(p_N, c + \int_{-\infty}^{\infty} V_{t-1}(\mathbf{p}, \rho)\, dF(\rho \mid \mathbf{p})\right).$$

For the one-period problem, let x be a continuation price given $\mathbf{p}$; let y be a strict stopping price given $\mathbf{p}$. (If x and y do not exist, there are trivial reservation prices.) Thus, from the one-period recursive expression,

$$E_{\rho|\mathbf{p},y} V_0(\mathbf{p}, y, \rho) > y - c, \tag{3}$$

$$E_{\rho|\mathbf{p},x} V_0(\mathbf{p}, x, \rho) \leqslant x - c. \tag{4}$$

But $V_0(\mathbf{p}, z) = z$ for any z, so (3) and (4) imply

$$E_{\rho|\mathbf{p},y}(\rho) - E_{\rho|\mathbf{p},x}(\rho) > y - x.$$

If $y > x$, this contradicts (1) which we have already established. Hence, $y \leqslant x$, and the optimal one-period policy is reservation price.

Finally, (2) follows from (1) since $V_0(\mathbf{p}, \rho) = \rho$.

This completes the proof for the one-period problem. Our induction hypotheses are that (2) holds and that the optimal policy is reservation price. Assume they hold up through t, and let us focus on the $(t+1)$-period problem.

For the t-period problem, let P be the reservation price given $\mathbf{p}$; let Q be the reservation price given $\mathbf{q}$.[5] Then,

$$\begin{aligned}
&E_{\rho\mid\mathbf{p}}V_t(\mathbf{p},\rho)-E_{\rho\mid\mathbf{q}}V_t(\mathbf{q},\rho)\\
&\quad=\int_{-\infty}^{P}\rho f(\rho\mid\mathbf{p})\,d\rho-\int_{-\infty}^{Q}\rho f(\rho\mid\mathbf{q})\,d\rho+\int_{P}^{\infty}V_t(\mathbf{p},\rho)f(\rho\mid\mathbf{p})\,d\rho\\
&\qquad-\int_{Q}^{\infty}V_t(\mathbf{q},\rho)\,f(\rho\mid\mathbf{q})\,d\rho.
\end{aligned}$$

Observe that $Q\leqslant P$ by induction hypothesis (2), and after some rearrangement and the appropriate addition and subtraction of like terms, the above becomes

$$\begin{aligned}
E_{\rho\mid\mathbf{p}}V_t(\mathbf{p},\rho)-E_{\rho\mid\mathbf{q}}V_t(\mathbf{q},\rho)=&\int_{-\infty}^{P}\rho[f(\rho\mid\mathbf{p})-f(\rho\mid\mathbf{q})]\,d\rho\\
&+\int_{P}^{\infty}V_t(\mathbf{p},\rho)[f(\rho\mid\mathbf{p})-f(\rho\mid\mathbf{q})]\,d\rho\\
&+\int_{Q}^{P}[\rho-V_t(\mathbf{q},\rho)]f(\rho\mid\mathbf{q})\,d\rho\\
&+\int_{P}^{\infty}[V_t(\mathbf{p},\rho)-V_t(\mathbf{q},\rho)]f(\rho\mid\mathbf{q})\,d\rho.
\end{aligned}\tag{5}$$

After integration by parts, the first two terms of (5) become

$$\begin{aligned}
&\int_{-\infty}^{P}[F(\rho\mid\mathbf{q})-F(\rho\mid\mathbf{p})]\,d\rho\\
&\quad+\int_{P}^{\infty}\partial/\partial\rho[V_t(\mathbf{p},\rho)][F(\rho\mid\mathbf{q})-F(\rho\mid\mathbf{p})]\,d\rho.
\end{aligned}\tag{6}$$

The second term is being integrated only over the region of continuation given $\mathbf{p}$; thus by the induction hypothesis

$$\partial/\partial\rho[V_t(\mathbf{p},\rho)]\leqslant\frac{1}{N+1}.$$

Hence, (6)

$$\begin{aligned}
&\leqslant\frac{1}{N+1}[E_{\rho\mid\mathbf{p}}(\rho)-E_{\rho\mid\mathbf{q}}(\rho)]+\frac{N}{N+1}\int_{-\infty}^{P}[F(\rho\mid\mathbf{q})-F(\rho\mid\mathbf{p})]\,d\rho\\
&\leqslant\frac{\Delta}{N(N+1)}+\frac{N}{N+1}[G(P\mid\mathbf{q})-G(P\mid\mathbf{p})].
\end{aligned}\tag{7}$$

---

[5] It may well be that P and Q depend on t, but this will be of no concern to us here.

The last two terms of (5) become, after integration by parts,

$$= [P - V_t(\mathbf{q}, P)]\, F(P \mid \mathbf{q}) - \int_Q^P \left[1 - \frac{\partial V_t(\mathbf{q}, \rho)}{\partial \rho}\right] F(\rho \mid \mathbf{q})\, d\rho$$
$$+ \int_P^\infty [V_t(\mathbf{p}, \rho) - V_t(\mathbf{q}, \rho)] f(\rho \mid \mathbf{q})\, d\rho. \tag{8}$$

Now note that we can fix P. Given this, observe from Fig. 1 that the maximal value of $E_{\rho \mid \mathbf{p}} V_t(\mathbf{p}, \rho) - E_{\rho \mid \mathbf{q}} V_t(\mathbf{q}, \rho)$ for a given $d = V_t(\mathbf{p}, P) - V_t(\mathbf{q}, P)$ occurs when $P - Q$ is as large as possible and when

$$\frac{\partial V_t(\mathbf{q}, \rho)}{\partial \rho} = \frac{1}{N+1} \qquad \text{for } Q \leqslant \rho \leqslant P.$$

In this case,

$$Q = P - \frac{N+1}{N} d.$$

Hence, (8) is

$$\leqslant dF(P \mid \mathbf{q}) - \int_{P-(N+1/N)d}^P \left[1 - \frac{1}{N+1}\right] F(\rho \mid \mathbf{q})\, d\rho$$
$$+ \int_P^\infty [V_t(\mathbf{p}, \rho) - V_t(\mathbf{q}, \rho)] f(\rho \mid \mathbf{q})\, d\rho. \tag{9}$$

By differentiating this expression with respect to d and noting from the diagram of Fig. 1 that $V_t(\mathbf{p}, \rho) - V_t(\mathbf{q}, \rho)$ increases as d increases, we see that the "worst" value of d [i.e., that value which maximizes (9)] is its maximal inductive value $\Delta/(N+1)$. So (9)

$$\leqslant \frac{\Delta}{N+1} F(P \mid \mathbf{q}) - \frac{N}{N+1} \int_{P-\Delta/N}^P F(\rho \mid \mathbf{q})\, d\rho$$
$$+ \int_P^\infty [V_t(\mathbf{p}, \rho) - V_t(\mathbf{q}, \rho)] f(\rho \mid \mathbf{q})\, d\rho$$
$$\leqslant \frac{\Delta}{N+1} F(P \mid \mathbf{q}) + \frac{N}{N+1} \left[G\left(P - \frac{\Delta}{N} \mid \mathbf{q}\right) - G(P \mid \mathbf{q})\right]$$
$$+ \int_P^\infty \frac{\Delta}{N+1} f(\rho \mid \mathbf{q})\, d\rho$$

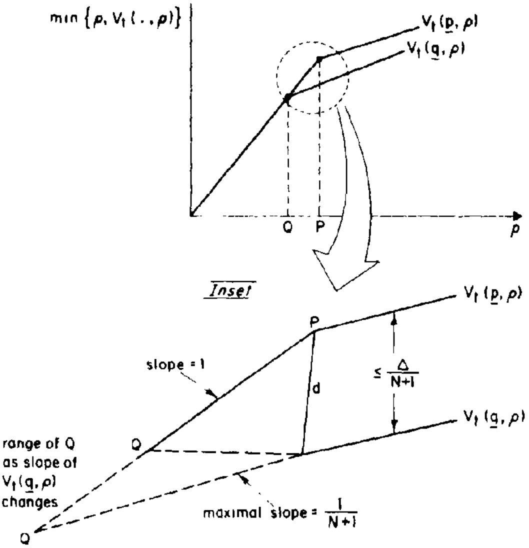


FIG. 1. The relationship between P and Q.

by induction, this

$$= \frac{\Delta}{N+1} + \frac{\Delta}{N+1}\left[G\left(P - \frac{\Delta}{N} \mid \mathbf{q}\right) - G(P \mid \mathbf{q})\right].$$

Combining this with (7), we have from (5) that

$$\begin{aligned} &E_{\rho|\mathbf{p}} V_t(\mathbf{p}, \rho) - E_{\rho|\mathbf{q}} V_t(\mathbf{q}, \rho) \\ &\leqslant \frac{\Delta}{N+1} + \frac{\Delta}{N(N+1)} + \frac{N}{N+1}\left[G\left(P - \frac{\Delta}{N} \mid \mathbf{q}\right) - G(P \mid \mathbf{p})\right] \\ &\leqslant \frac{\Delta}{N} \end{aligned}$$

from the hypotheses of the theorem. The non-negativity of (2) follows trivially by induction, and, thus (2) is established for the $(t + 1)$-period problem. The proof of the reservation price property now follows exactly as in the one-period case by use of (2).

The first condition requires that $\tilde{x}|\mathbf{p}$ stochastically dominate $\tilde{x}|\mathbf{q}$, or, loosely, that with higher price observations, one tends to believe that future observations will be higher. If this condition holds, then

$$G(x \mid \mathbf{p}) \leqslant G(x \mid \mathbf{q}),$$

but the second condition of the theorem says, "not too much less" since it must hold that

$$G\left(x+\frac{\Delta}{N} \mid \mathbf{p}\right) \geqslant G(x \mid \mathbf{q}).$$

If the distribution of the next price given past observations $\mathbf{p}$ is merely a shifted version of the distribution of the next price given history $\mathbf{q}$, and if that shift is less than Δ/N in magnitude, the conditions of Theorem 2 will hold. Hence, one would expect that distributions with unknown mean-related parameters might be good candidates for reservation price optimal policies.

We present two examples—the first is of this type.

Example 1: *Normal Distribution with Unknown Mean*

If we take $\tilde{x} \mid \mu \sim N(\mu, 1)$ with $\tilde{\mu} \sim N(\mu_0, 1/\tau)$, then, after observing $\mathbf{p}$, the posterior on $\tilde{\mu}$ is $N(\bar{\mu}, 1/(\tau+N))$ with

$$\bar{\mu}=\frac{\tau\mu_0+\sum p_i}{\tau+N}$$

and $F(x \mid \mathbf{p})$ is $N(\bar{\mu}, 1+1/(\tau+N))$. Note that since $F(x \mid \mathbf{q})$ is $N(\bar{\mu}-\Delta/(\tau+N), 1+1/(\tau+N))$,

$$F(x \mid \mathbf{q})=F\left(x+\frac{\Delta}{\tau+N} \mid \mathbf{p}\right),$$

so the first condition of the theorem holds. Also, the second condition is satisfied since

$$G\left(x+\frac{\Delta}{N} \mid \mathbf{p}\right) \geqslant G\left(x+\frac{\Delta}{\tau+N} \mid \mathbf{p}\right)=\int_{-\infty}^{x+\Delta/(\tau+N)} F(\rho \mid \mathbf{p})\, d\rho$$

$$=\int_{-\infty}^{x} F\left(\rho+\frac{\Delta}{\tau+N} \mid \mathbf{p}\right) d\rho=G(x \mid \mathbf{q}).$$

Example 2: *Exponential with Unknown Mean—Exponential Prior on Parameter* λ

We take the forecasting distribution of the next price to be $F(x \mid \lambda)=1-e^{-\lambda x}$, and the prior on λ to have initial density $e^{-\alpha\lambda}$. It may be shown, after noting that

$$\int_0^{\infty} \lambda^N e^{-\gamma\lambda}\, d\lambda=\frac{N!}{\gamma^{N+1}}, \tag{10}$$

that the posterior on λ after observing prices $\mathbf{p} = p_1, p_2, \ldots, p_N$ is

$$h(\lambda \mid \mathbf{p}) = \frac{A^{N+1}}{N!} \lambda^N e^{-A\lambda},$$

where

$$A = \sum_{i=1}^{N} p_i + \alpha.$$

Hence,

$$\begin{aligned} F(x \mid \mathbf{p}) &= \int_0^x \frac{A^{N+1}}{N!} \lambda^N e^{-A\lambda} (1 - e^{-A\lambda})(1 - e^{-\lambda x})\, d\lambda \\ &= 1 - \left(\frac{A}{A+x}\right)^{N+1}. \end{aligned}$$

Thus, for $\mathbf{p} \geqslant \mathbf{q}$, $F(x \mid \mathbf{p}) \leqslant F(x \mid \mathbf{q})$.

To verify the second condition, observe that

$$G(x \mid \lambda) = \int_0^x F(\rho \mid \lambda)\, d\rho = x - \frac{1}{\lambda} + \frac{e^{-\lambda x}}{\lambda}.$$

Hence,

$$\begin{aligned} &G\left(x + \frac{\Delta}{N} \mid \mathbf{p}\right) - G(x \mid \mathbf{q}) \\ &\quad = \left\{\frac{\Delta}{N} - \left[E_{\lambda \mid \mathbf{p}}\left(\frac{1}{\lambda}\right) - E_{\lambda \mid \mathbf{q}}\left(\frac{1}{\lambda}\right)\right]\right\} \\ &\qquad + \left\{E_{\lambda \mid \mathbf{p}}\left(\frac{e^{-\lambda(x+\Delta/N)}}{\lambda}\right) - E_{\lambda \mid \mathbf{q}}\left(\frac{e^{-\lambda x}}{\lambda}\right)\right\} \\ &\quad \sim \frac{A_p^{N+1}}{(A_p + x + \Delta/N)^N} - \frac{A_q^{N+1}}{(A_q + x)^N} \end{aligned}$$

by a double application of (10).

Let I be the index of the component which differs in $\mathbf{p}$ and $\mathbf{q}$. Then we can write this as

$$\begin{aligned} &\frac{(\alpha + \sum_{i \neq I} p_i + p_I)^{N+1}}{(\alpha + \sum_{i \neq I} p_i + x - q_I/N + ((N+1)/N)\, p_I)^N} \\ &\qquad - \frac{(\alpha + \sum_{i \neq I} p_i + q_I)^{N+1}}{(\alpha + \sum_{i \neq I} p_i + x - q_I/N + ((N+1)/N)\, q_I)^N} \\ &= \frac{(C_1 + p_I)^{N+1}}{(C_2 + ((N+1)/N)\, p_I)^N} - \frac{(C_1 + q_I)^{N+1}}{(C_2 + ((N+1)/N)\, q_I)^N} \end{aligned}$$

for appropriate constants C_1 and C_2.

By differentiating the function $(C_1 + x)^{N+1}/(C_2 + ((N+1)/N)\,x)^N$, we can demonstrate the non-negativity of the above expression.

Some comparison of Theorem 2 with Rothschild's [15] result for the Dirichlet distribution is worthwhile here. The Dirichlet, in general, does not satisfy the second condition of Theorem 2. The difference between $G(x \mid \mathbf{q})$ and $G(x \mid \mathbf{p})$ can be as large as $\Delta/(N+m)$; where m is the number of prior equivalent observations. The line of proof that Rothschild uses capitalizes on the aspects of the Dirichlet-multinomial. Theorem 2, as noted above, is often applicable to distributions with known functional forms but unknown means. The Dirichlet-multinomial is not of this type. Theorem 2 is thus complementary to Rothschild's result; the two results together cover a range of distributions.

VI. FINAL COMMENTS

This paper has derived conditions guaranteeing optimal myopic policies for recall allowed problems and optimal reservation policies for both classes of problems. These optimal policy properties have been used extensively in economic models. For example, Lippman and McCall have modelled the job search process [8, 9]; Masson relates price search policies to racial price discrimination [10]; Pratt *et al.* [13] have described empirical evidence on consumer buying behavior and its relation to theoretical results.

This article should be of use to economic modelers in assessing whether the structure of the price uncertainty appropriate to their research permits these useful regularity conditions to hold.

APPENDIX: VERIFICATION THAT EXPONENTIAL WITH UNKNOWN ORIGIN IS CHARACTERIZED BY A MYOPIC OPTIMAL POLICY WITH INCREASING RESERVATION PRICES

The validity of the example follows from the following generalization:

COROLLARY 2. *Suppose*

(1) *Prices are distributed according to* $F(p - t)$, t *being an unknown origin.*

(2) *The posterior for* t, $h(t \mid N, p_N^*)$, *depends on* N *and* p_N^* *only.*

(3) $h(t \mid N, p_N^*)$ *has monotone likelihood ratio* (MLR) *in* N *for fixed* p_N^*. (*A density* $f(x \mid \theta)$ *is said to have* MLR *in* θ *if for* $\theta_1 > \theta_2$, $f(x \mid \theta_1)/f(x \mid \theta_2)$ *is non-decreasing in* x.)

(4) $h(p_N^* - u \mid N, p_N^*)$ *has* MLR *in* p_N^* *for fixed* N.

Then the conditions of Corollary 1 *are satisfied and the optimal policy is myopic with increasing reservation prices.*

Proof.

$$
\begin{aligned}
G(p_N^* \mid \mathbf{p}) &= \int_{-\infty}^{p_N^*} F(\rho \mid \mathbf{p})\, d\rho \\
&= \int_{-\infty}^{p_N^*} \int_{-\infty}^{\rho} F(\rho - t)\, h(t \mid N, p_N^*)\, dt\, d\rho \\
&= \int_{-\infty}^{p_N^*} G(p_N^* - t)\, h(t \mid N, p_N^*)\, dt.
\end{aligned}
$$

Observe that $G(p_N^* \mid \mathbf{p})$ depends only on N and p_N^* so we need only show the monotonicity of G in N and p_N^*.

It can be shown that if a density $f(x \mid \theta)$ has MLR in θ, γ is a function which crosses zero at most once, then

$$\int_{-\infty}^{\infty} \gamma(x)\, f(x \mid \theta)\, dx$$

crosses zero at most once in θ and in the same direction as γ. If γ is monotonic, then by defining

$$\gamma_c(x) = \gamma(x) - c,$$

we see that

$$\int_{-\infty}^{\infty} \gamma_c(x)\, f(x \mid \theta)\, dx$$

crosses c at most once for any c, and, hence, is monotonic in θ in the same direction as γ.

Since $G(p_N^* - t)$ is non-increasing in t has the above form and hence $G(p_N^* \mid \mathbf{p})$ is non-increasing in N.

Furthermore,

$$
\begin{aligned}
G(p_N^* \mid \mathbf{p}) &= \int_{-\infty}^{p_N^*} G(p_N^* - t)\, h(t \mid N, p_N^*)\, dt \\
&= \int_{0}^{\infty} G(u)\, h(p_N^* - u \mid N, p_N^*)\, du
\end{aligned}
$$

From condition (4) of the corollary, this has the required form and thus G is monotonic in p_N^*.

For the exponential, we can, without loss of generality take

$$F(p-t)=1-e^{-(p-t)}.$$

Denote by I_A the indicator function of the event A, and let $r(t)$, with $\log r(t)$ concave, be the prior on the unknown origin t. It follows that

$$h(t\,|\,N,p_N^*)=Kr(t)\,e^{Nt}I_{(t<p_N^*)},$$

where K is a normalizing constant (which depends on p_N^*). For fixed ξ,

$$\frac{h(t\,|\,M,\xi)}{h(t\,|\,N,\xi)}=\text{constant}\times e^{(M-N)t}$$

and condition (3) follows. For a fixed N and $y\geqslant x$,

$$\frac{h(y-u\,|\,N,y)}{h(x-u\,|\,N,x)}=C\frac{r(y-u)}{r(x-u)}\qquad \text{for } u\geqslant 0,$$

where C is a constant depending on x and y.

The monotonicity of $r(y-u)/r(x-u)$ or equivalently $\log r(y-u)-\log r(x-u)$ follows directly from the log-concavity of $r(t)$.

ACKNOWLEDGEMENT

We are grateful to John Pratt for kindling our interest in this problem and for providing many useful suggestions throughout the course of the research, and to David Butler for reviewing the technical details.

REFERENCES

1. M. ABDEL-HAMEED, Optimality of the one step look-ahead stopping times, *J. Appl. Prob.* **14** (1977), 162–169.
2. C. ALBRIGHT, A Bayesian approach to a generalized house selling problem, *Manage. Sci.* **24**, No. 2 (1977), 432–440.
3. L. BREIMAN, Stopping rule problems, *in* "Applied Combinatorial Mathematics" (E, F. Beckenbach, Ed.), Wiley, New York, 1904.
4. Y. S. CHOW AND H. ROBBINS, A Martingale system theory and applications, *in* "Proceedings, Fourth Berkeley Symposium, Berkeley, Cal.," Univ. of California Press, Berkely, 1961.
5. M. DEGROOT, "Optimal Statistical Decisions," McGraw–Hill, New York, 1970.
6. R. HAYES, Optimal strategies for divestitures, *Operations Res.* **17** (1970), 292–310.
7. M. KOHN AND S. SHAVELL, The theory of search, *J. Econ. Theory* **9**, No. 2 (1974), 93–123.
8. S. LIPPMAN AND J. MCCALL, Job search in a dynamic economy, *J. Econ. Theory* **12**, No. 3 (1976), 365–390.

9. S. Lippman and J. McCall, The economics of job search: A survey, *Econ. Inquiry* **14**, No. 2 (1976), 155–189 and No. 3 (1976), 347–368.
10. R. T. Masson, Costs of search and racial price discrimination, *Western Econ. J.* **11**, No. 2 (1973), 167–186.
11. J. McCall, The economics of information and optimal stopping rules, *J. Nusiness* **38** (1965), 300–317.
12. J. McCall, Economics of information and job search, *Quart. J. Econ.* **84** (1970) 113–126.
13. J. Pratt, D. Wise, and R. Zeckhauser, Price differences in almost competitiv markets, *Quart. J. Economics* (May 1979).
14. M. Rothschild, Models of market organization with imperfect information, *J. Pol. Econ* **81** (1973), 1283–1308.
15. M. Rothschild, Searching for the lowest price when the distribution of prices i unknown, *J. Pol. Econ.* **8**, No. 4 (1974), 689–711.
16. G. Stigler, The economics of information, *J. Pol. Econ.* **69** (1061), 213–225.
17. L. Telser, Searching for the lowest price, *Amer. Econ. Rev.* **63**, No. 2 (1973), 40–49.
18. J. Yahav, On optimal stopping, *Ann. Math. Statist.* **37** (1966), 30–35.

Dynamic Games of Innovation

JENNIFER F. REINGANUM

California Institute of Technology, Pasadena, California 91125

Received July 18, 1979; revised July 22, 1980

I. INTRODUCTION

The purpose of this paper is the investigation of issues related to industrial research and development. In particular, we will focus on situations in which two firms are rivals in developing a new process or device. There is often a distinct advantage to being the first to produce a new product or implement a new technology. This advantage may be conferred, for example, by the award of a patent. But since only the first to succeed realizes this advantage, each firm's profits will depend upon the research efforts of its rivals. This interaction of profit functions naturally suggests a game theoretic approach.

In what follows, we will develop a theory of optimal resource allocation to research and development under the assumption of uncertain technical advance and in the presence of game-playing rivals. We will focus initially upon the development and properties of noncooperative and cooperative equilibria. Later sections examine the social optimality of the market (game) equilibrium and the implications which result for patent, tax/subsidy and anti-trust policy. While we find that the Nash equilibrium and the socially optimal rates of investment in *R & D* do not coincide, it is not clear (a priori) in which direction the bias acts. Finally, the public good aspect of knowledge production is incorporated and its effect on equilibrium investment in *R & D* is determined.

II. BACKGROUND

By the term "research and development with rivalry" we mean a situation in which firms are competing for the profits which accrue with the introduction of a particular innovation. In particular, we will focus on situations in which two firms are rivals in developing a new process or device. This rivalry may be enhanced or diminished by the availability of patent protection and the inability to keep research findings secret. Throughout this paper we will consider only the case of perfect patent

0022-0531/81/040021-21$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved.

protection. That is, the first firm to perfect the innovation is granted a legal monopoly over the disposition of the innovation, while imitation is forbidden. For an analysis of imperfect patent protection and its impact on equilibrium investment in *R* & *D*, see Reinganum [10].

An extremely important feature of innovation production is uncertainty regarding the timing of successful innovation. Because of this uncertainty, firms must make contingency plans for investing in *R* & *D* over some (possibly protracted) planning horizon. Since a firm may be pre-empted by a rival at any time, it is imperative that the firm be able to adjust its expenditures on *R* & *D* at each date of the planning horizon. Thus a reasonable model of induced technical change must address the dynamic nature of contingency planning.

In the literature on research and development with rivalry, basically two alternative approaches have been used to analyze the situation outlined above:

(1) A game theoretic analysis (e.g., Scherer [13]; Loury [9]; Lee and Wilde [8]; Dasgupta and Stiglitz [2, 3]), which is static in the sense that each firm chooses its investment strategy from a subset of R^+.

(2) A decision theoretic analysis by Kamien and Schwartz [5–7], which is dynamic in the sense that each firm chooses a function, a time path of spending on *R* & *D*.

The model developed in this paper represents a synthesis of the two approaches described above, combining what we believe to be the best elements of each method. Like the game theoretic analysis, we will model the rivals as strategic agents, able not only to calculate their own best responses to rivals' given strategies, but their rivals' best responses and the ensuing Nash equilibrium as well. Like the decision theoretic analysis, we allow for dynamic or contingency planning. That is, the players will choose a strategy which may be expressed as a function of time. The resulting analytical technique is that of differential games or games played in continuous time.

III. The Basic Model

We assume that there are two identical firms (players) engaged in a race for a particular technological breakthrough—an invention or process. While the game easily generalizes to *n* identical firms, in this paper we will not be explicitly concerned with the effects of the number of rivals on equilibrium investment in *R* & *D*. However, this question is addressed in Reinganum [10].

Definition 1. We say that player *i* *succeeds* if *i* perfects the innovation.

DEFINITION 2. We say that player i *wins* if i succeeds before the rival does.

If firm i wins the race, then i receives a patent whose worth is the present value (or expected present value if the reward stream is stochastic) of the stream of rents accruing to a monopolist in the market for the innovation. If firm i loses or simply fails to succeed, then i gains nothing. In both cases, i forfeits its research investment to date.

We assume that each player accumulates knowledge relevant to the particular project via the expenditure of resources on research activity or knowledge acquisition. This knowledge accumulates deterministically—that is, research activity results in a steady, direct increase in the stock of relevant knowledge. No allowance is made for decreases or for random changes in the knowledge stock. Innovation, however, is assumed to be stochastic. The time of successful completion of the device or process is a random variable whose distribution is under the partial control of the firm. That is, by increasing its stock of relevant knowledge, i can increase its likelihood of success by a given date but it cannot choose its exact date of completion.

N1. T is the calendar date of Doomsday. This may be regarded as a time at which the invention becomes completely obsolete due to some discontinuity in technological advance. Alternatively, it may represent a date at which time the firms abandon the project entirely if they haven't perfected the innovation. That is, if a firm fails to succeed by T, it becomes convinced that the innovation is not feasible.

N2. r is the common discount rate.

N3. $P(t)$ is the present value of a patent obtained at t. For simplicity, we will assume that this discounted patent value is a constant P. Because the present value of the patent is the same regardless of the date of success, this assumption eliminates the influence of patent value growth or decline on investment decisions. It thus serves to focus incentives to invest upon the possibility of the rival's pre-emptive success.

N4. t_i is the (random) time at which firm i succeeds.

N5. $\mu_i(t)$ is the rate at which firm i acquires knowledge at t; $\mu_i(t)$ represents the additional relevant knowledge that firm i chooses to acquire at time t. We further assume that $\mu_i(t) \in [0, B]$, where B is a positive constant. That is, there is an exogenous physical limit upon the rate at which knowledge can be acquired.

N6. $e^{-rt}c_i(\mu_i)$ is the discounted cost of additional knowledge μ_i acquired at t; we assume throughout that $c_i(\mu_i) = (1/2)(\mu_i)^2$, although the generalization to any constant elasticity cost function is straightforward.

discounting the flow costs at rate r, we can write firm 1's payoff for any strategy pair (u_1, u_2) as follows:

$$J^1(u_1, u_2) = \int_0^T \{Pe^{-\lambda(z_1(t)+z_2(t))}\lambda\mu_1(t) - e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)(\mu_1(t))^2\}\, dt,$$

where $\mu_i(t) = u_i(t, z(t))$ and $z(t)$ solves $\dot{z} = u(t, z)$ with $z(0) = 0$. We use the abbreviated (u_1, u_2) to represent the entire strategy pair $(u_1(t, z), u_2(t, z))$ over all of $[0, T] \times Z_1 \times Z_2$.

It will prove more convenient to integrate the first term of the payoff by parts and deal with the resulting equivalent payoff:

$$\begin{aligned} J^1(u_1, u_2) = \int_0^T & \{P(1 - e^{-\lambda z_1(t)})e^{-\lambda z_2(t)}\lambda\mu_2(t) \\ & - e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)(\mu_1(t))^2\}\, dt \\ & + P(1 - e^{-\lambda z_1(T)})\, e^{-\lambda z_2(T)}. \end{aligned}$$

The analog for player 2 is obvious.

IV. THE NONCOOPERATIVE GAME

In order to propose a method of solution, we must first decide upon the players' behavioral attitudes. We first assume that each firm maximizes its own payoff, taking its rival's strategy as given. This assumption results in the Nash equilibrium concept.

DEFINITION 8. A strategy pair (u_1^*, u_2^*) is a *Nash equilibrium* if

(1) $u_i^*(t, z) \in \mathscr{U}_i$, $i = 1, 2$;
(2) $J^1(u_1^*, u_2^*) \geqslant J^1(u_1, u_2^*)$ for all $u_1 \in \mathscr{U}_1$;
(3) $J^2(u_1^*, u_2^*) \geqslant J^2(u_1^*, u_2)$ for all $u_2 \in \mathscr{U}_2$.

For the moment, we will assume that a Nash equilibrium exists and we will attempt to characterize it. Let

$$\begin{aligned} V^i(s, y) \equiv \int_s^T & \{(1 - e^{-\lambda z_i(t)})e^{-\lambda z_j(t)}P\lambda\mu_j^* - e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)(u_i^*)^2\}\, dt \\ & + P(1 - e^{-\lambda z_i(T)})\, e^{-\lambda z_j(T)}, \end{aligned}$$

where $y = z(s)$.

Then from Theorem 8.2.2 (Friedman, [4]), we know that (u_1^*, u_2^*) must satisfy a system of Bellman equations:

$$V_t^i(t,z) + \max_{u_i(t,z)\in[0,B]} [V_{z_i}^i(t,z)\,u_i(t,z) + V_{z_j}^i(t,z)\,u_j^*(t,z)$$

$$+ P(1-e^{-\lambda z_i})e^{-\lambda z_j}\lambda u_j^*(t,z) - e^{-rt}e^{-\lambda(z_1+z_2)}(1/2)(u_i(t,z))^2] = 0. \quad (2)$$

Then by definition of $V^i(s, y)$, it must be that

$$V^i(T, z(T)) = Pe^{-\lambda z_j(T)} - Pe^{-\lambda(z_1(T)+z_2(T))}. \quad (3)$$

The candidates for a Nash equilibrium can be obtained by performing the indicated maximization. Since the left-hand side of Eq. (2) is strictly concave in u_1, the necessary conditions for a maximum are also sufficient.

$$\begin{aligned} V_{z_i}^i(t,z) - e^{-rt}e^{-\lambda(z_1+z_2)}u_i(t,z) &> 0\ \forall u_i(t,z)\in[0,B] \Rightarrow u_i^*(t,z) = B.\\ &= 0 \Rightarrow u_i^*(t,z) = V_{z_i}^i(t,z)\,e^{rt}e^{\lambda(z_1+z_2)}.\\ &< 0\ \forall u_i(t,z)\in[0,B] \Rightarrow u_i^*(t,z) = 0. \end{aligned}$$

These conditions state that if $u_i^*(t,z)$ is to be within the interval $(0, B)$, then the discounted within-period expected marginal cost $e^{-rt}e^{-\lambda(z_1+z_2)}\,u_i^*(t,z)$ must be exactly offset by the additional expected profit to the firm generated by that cost—$V_{z_i}^i(t,z)$. If the additional expected profit is below the marginal cost of any positive level of R & D, then $u_i^*(t,z) = 0$. If the additional expected profit exceeds even the marginal cost of advancing at the maximal rate B, then the firm undertakes an all-out push for completion of the project, $u_i^*(t,z) = B$.

Unfortunately, we can say very little a priori about the exact functional form of the value functions and their dependence upon the parameters λ, r and T. Thus it is difficult to give a full description of the equilibrium strategies.

Suppose we look for interior solutions, where $u_i^*(t,z)\in(0,B)$ for all $(t,z)\in[0,T]\times Z_1\times Z_2$. If we substitute the expressions $u_i^\circ = V_{z_i}^i e^{rt}e^{\lambda(z_1+z_2)}$ back into the Bellman system (2), and if we can solve the resulting system of Hamilton–Jacobi equations for the value functions $V^1(t,z)$ and $V^2(t,z)$, then we can determine the Nash equilibrium (candidate) strategies via the substitution

$$u_i^*(t,z) = u_i^\circ(t, z, V_{z_i}^i(t,z)).$$

Substituting the expressions (u_1°, u_2°) into the system (2) and simplifying yields:

$$V_t^i + (V_{z_i}^i)^2(1/2)e^{rt}e^{\lambda(z_1+z_2)} + V_{z_j}^i V_{z_j}^j e^{rt}e^{\lambda(z_1+z_2)} + V_{z_j}^i P e^{rt}\lambda(e^{\lambda z_i} - 1) = 0 \quad (4)$$

with terminal conditions (3).

discounting the flow costs at rate r, we can write firm 1's payoff for any strategy pair (u_1, u_2) as follows:

$$J^1(u_1, u_2) = \int_0^T \{Pe^{-\lambda(z_1(t)+z_2(t))}\lambda\mu_1(t) - e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)(\mu_1(t))^2\}\,dt,$$

where $\mu_i(t) = u_i(t, z(t))$ and $z(t)$ solves $\dot{z} = u(t, z)$ with $z(0) = 0$. We use the abbreviated (u_1, u_2) to represent the entire strategy pair $(u_1(t, z), u_2(t, z))$ over all of $[0, T] \times Z_1 \times Z_2$.

It will prove more convenient to integrate the first term of the payoff by parts and deal with the resulting equivalent payoff:

$$\begin{aligned} J^1(u_1, u_2) = \int_0^T \{&P(1 - e^{-\lambda z_1(t)})e^{-\lambda z_2(t)}\lambda\mu_2(t) \\ &- e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)(\mu_1(t))^2\}\,dt \\ &+ P(1 - e^{-\lambda z_1(T)})\,e^{-\lambda z_2(T)}. \end{aligned}$$

The analog for player 2 is obvious.

IV. THE NONCOOPERATIVE GAME

In order to propose a method of solution, we must first decide upon the players' behavioral attitudes. We first assume that each firm maximizes its own payoff, taking its rival's strategy as given. This assumption results in the Nash equilibrium concept.

DEFINITION 8. A strategy pair (u_1^*, u_2^*) is a *Nash equilibrium* if

(1) $u_i^*(t, z) \in \mathscr{U}_i$, $i = 1, 2$;
(2) $J^1(u_1^*, u_2^*) \geqslant J^1(u_1, u_2^*)$ for all $u_1 \in \mathscr{U}_1$;
(3) $J^2(u_1^*, u_2^*) \geqslant J^2(u_1^*, u_2)$ for all $u_2 \in \mathscr{U}_2$.

For the moment, we will assume that a Nash equilibrium exists and we will attempt to characterize it. Let

$$\begin{aligned} V^i(s, y) \equiv \int_s^T \{&(1 - e^{-\lambda z_i(t)})e^{-\lambda z_j(t)}P\lambda\mu_j^* - e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)(u_i^*)^2\}\,dt \\ &+ P(1 - e^{-\lambda z_i(T)})\,e^{-\lambda z_j(T)}, \end{aligned}$$

where $y = z(s)$.

Then from Theorem 8.2.2 (Friedman, [4]), we know that (u_1^*, u_2^*) must satisfy a system of Bellman equations:

$$V_t^i(t,z) + \max_{u_i(t,z)\in[0,B]} [V_{z_i}^i(t,z)\,u_i(t,z) + V_{z_j}^i(t,z)\,u_j^*(t,z)$$

$$+ P(1-e^{-\lambda z_i})e^{-\lambda z_j}\lambda u_j^*(t,z) - e^{-rt}e^{-\lambda(z_1+z_2)}(1/2)(u_i(t,z))^2] = 0. \qquad (2)$$

Then by definition of $V^i(s,y)$, it must be that

$$V^i(T, z(T)) = Pe^{-\lambda z_j(T)} - Pe^{-\lambda(z_1(T)+z_2(T))}. \qquad (3)$$

The candidates for a Nash equilibrium can be obtained by performing the indicated maximization. Since the left-hand side of Eq. (2) is strictly concave in u_1, the necessary conditions for a maximum are also sufficient.

$$\begin{aligned} V_{z_i}^i(t,z) - e^{-rt}e^{-\lambda(z_1+z_2)}u_i(t,z) &> 0\ \forall u_i(t,z)\in[0,B] \Rightarrow u_i^*(t,z) = B. \\ &= 0 \Rightarrow u_i^*(t,z) = V_{z_i}^i(t,z)\,e^{rt}e^{\lambda(z_1+z_2)}. \\ &< 0\ \forall u_i(t,z)\in[0,B] \Rightarrow u_i^*(t,z) = 0. \end{aligned}$$

These conditions state that if $u_i^*(t,z)$ is to be within the interval $(0,B)$, then the discounted within-period expected marginal cost $e^{-rt}e^{-\lambda(z_1+z_2)}u_i^*(t,z)$ must be exactly offset by the additional expected profit to the firm generated by that cost—$V_{z_i}^i(t,z)$. If the additional expected profit is below the marginal cost of any positive level of R & D, then $u_i^*(t,z) = 0$. If the additional expected profit exceeds even the marginal cost of advancing at the maximal rate B, then the firm undertakes an all-out push for completion of the project, $u_i^*(t,z) = B$.

Unfortunately, we can say very little a priori about the exact functional form of the value functions and their dependence upon the parameters λ, r and T. Thus it is difficult to give a full description of the equilibrium strategies.

Suppose we look for interior solutions, where $u_i^*(t,z)\in(0,B)$ for all $(t,z)\in[0,T]\times Z_1\times Z_2$. If we substitute the expressions $u_i^\circ = V_{z_i}^i e^{rt}e^{\lambda(z_1+z_2)}$ back into the Bellman system (2), and if we can solve the resulting system of Hamilton–Jacobi equations for the value functions $V^1(t,z)$ and $V^2(t,z)$, then we can determine the Nash equilibrium (candidate) strategies via the substitution

$$u_i^*(t,z) = u_i^\circ(t,z,V_{z_i}^i(t,z)).$$

Substituting the expressions (u_1°, u_2°) into the system (2) and simplifying yields:

$$V_t^i + (V_{z_i}^i)^2(1/2)e^{rt}e^{\lambda(z_1+z_2)} + V_{z_j}^i V_{z_j}^j e^{rt}e^{\lambda(z_1+z_2)} + V_{z_j}^j Pe^{rt}\lambda(e^{\lambda z_i} - 1) = 0 \qquad (4)$$

with terminal conditions (3).

We propose a solution of the form $V^i(t,z) = b(t)e^{-\lambda(z_1+z_2)} + k(t)e^{-\lambda z_j}$.

Since $V^1_{z_1}(t,z) = V^2_{z_2}(t,z) = -b(t)e^{-\lambda(z_1+z_2)}$, we can reduce the system of Hamilton–Jacobi equations to two independent partial differential equations. Then each can be solved separately.

Substituting the appropriate partial derivatives and collecting coefficients, Eq. (4) becomes

$$\begin{aligned} &e^{-\lambda(z_1+z_2)}[\dot b(t) + (b(t))^2\lambda^2 e^{rt}(3/2) + b(t)\,\lambda^2 P e^{rt}] \\ &\quad + e^{-\lambda z_j}[\dot k(t) + k(t)\,b(t)\,\lambda^2 e^{rt} - b(t)\,\lambda^2 P e^{rt}] = 0. \qquad (5) \end{aligned}$$

Since $e^{-\lambda(z_1+z_2)}$ and $e^{-\lambda z_j}$ are always positive, the coefficients $b(t)$ and $k(t)$ must solve the ordinary differential equations and boundary conditions

$$\dot b(t) + (b(t))^2\lambda^2 e^{rt}(3/2) + b(t)\,\lambda^2 P e^{rt} = 0, \qquad b(T) = -P. \qquad (6)$$

$$\dot k(t) + k(t)\,b(t)\,\lambda^2 e^{rt} - b(t)\,\lambda^2 P e^{rt} = 0, \qquad k(T) = P. \qquad (7)$$

Equation (6) can be reduced to a linear ordinary differential equation via the substitution $q(t) = -1/b(t)$ and thus has the unique solution

$$b(t; -P) = \frac{-2P}{3 - \exp\{P\lambda^2(e^{rt} - e^{rT})/r\}} < 0 \qquad \text{for all} \quad t \leqslant T.$$

The second coefficient $k(t)$ must solve (7) with the substitution $b(t) = b(t; -P)$. The unique solution to this nonhomogeneous linear ordinary differential equation satisfying $k(T) = P$ is $k(t; P) \equiv P$.

We can now write the value functions for Nash rivals:

$$V^i(t,z) = \frac{-2Pe^{-\lambda(z_1+z_2)}}{3 - \exp\{m(t)\}} + Pe^{-\lambda z_j},$$

where $m(t) = P\lambda^2(e^{rt} - e^{rT})/r$.

We can see that the marginal benefit to i of an increment to i's own knowledge stock is

$$V^i_{z_i}(t,z) = \frac{2P\lambda e^{-\lambda(z_1+z_2)}}{3 - \exp\{m(t)\}} > 0, \qquad \text{for all } (t,z).$$

An increment to the rival's knowledge stock has two conflicting effects. On the one hand, it reduces the firm's chances of obtaining the patent. On the other, it reduces the expected costs firm i will incur. Thus

$$V^i_{z_j}(t,z) = \frac{P\lambda e^{-\lambda z_j}[2e^{-\lambda z_i} - 3 + \exp\{m(t)\}]}{3 - \exp\{m(t)\}}.$$

Since $2e^{-\lambda z_i} < 2$ for $t > 0$ and $\exp\{m(t)\} < 1$ for $t < T$, the marginal loss in expected patent benefits exceeds the marginal benifit of reduced expected costs. That is, $V^i_{z_j} < 0$ for all $(t, z) \in [0, T] \times Z_1 \times Z_2$.

We can now state

PROPOSITION 1. *Suppose that* $B \geqslant P\lambda e^{rT}$. *Then*

$$\text{(a)} \quad u_i^*(t, z) = \frac{2P\lambda e^{rt}}{3 - \exp\{P\lambda^2(e^{rt} - e^{rT})/r\}}, \qquad i = 1, 2,$$

is the unique Nash equilibrium (in pure strategies) in $\mathscr{U}_1 \times \mathscr{U}_2$;

(b) *The Nash equilibrium allocation is*

$$J^i(u_1^*, u_2^*) = P - \frac{2P}{3 - \exp\{P\lambda^2(1 - e^{rT})/r\}}, \qquad i = 1, 2.$$

Proof. (a) Clearly $(u_1^*, u_2^*) \in \mathscr{U}_1 \times \mathscr{U}_2$. We presented a continuously differentiable solution $(V^1(t, z), V^2(t, z))$ to the system (4) and terminal conditions (3) above. Then by construction, $V^i(t, z)$ and $u_i^*(t, z) = V^i_{z_i}(t, z)e^{rt}e^{\lambda(z_1 + z_2)}$ satisfy the hypotheses of sufficiency Theorem 1 (Stalford and Leitmann [15]). Therefore

$$u_i^*(t, z) = V^i_{z_i}(t, z)e^{rt}e^{\lambda(z_1 + z_2)} = \frac{2P\lambda e^{rt}}{3 - \exp\{P\lambda^2(e^{rt} - e^{rT})/r\}}$$

for $i = 1, 2$ is a Nash equilibrium in pure strategies. Since the left-hand side of system (4) is of class C^∞ in (t, z, V_z) and since the terminal functions in (3) are of class C^∞ in z, by Theorem 18.1 (Bernstein [1]), the solution $(V^1(t, z), V^2(t, z))$ is unique. Thus $(u_1^*(t, z), u_2^*(t, z))$ as defined above is the unique Nash equilibrium in $\mathscr{U}_1 \times \mathscr{U}_2$.

(b) The value of profits obtained by playing the Nash equilibrium strategies are

$$J^i(u_1^*, u_2^*) \equiv V^i(0, 0) = P - \frac{2P}{3 - \exp\{P\lambda^2(1 - e^{rT})/r\}}, \qquad i = 1, 2.$$

Due to the "memorylessness" property of the exponential distribution, the control functions are the strategies themselves. That is, the equilibrium strategies are functions of time only. They are strictly positive on $[0, T]$ and, for $B \geqslant P\lambda e^{rT}$, $u_i^*(t, z) \leqslant B$ for all t in $[0, T]$. Thus for this admittedly special case, a unique Nash equilibrium (in pure strategies) exists for all (t, z) in $[0, T] \times Z_1 \times Z_2$.

Direct differentiation yields:

$\partial u_i^*/\partial P > 0$; as the discounted patent value increases, so does the optimal rate of knowledge acquisition.

$\partial u_i^*/\partial T < 0$; as the planning horizon increases, firms spread out their R & D, acquiring knowledge at a lower rate.

An increase in the exponential parameter λ results in accelerated development. To see this, note that

$$\partial u_i^*/\partial \lambda = \frac{2Pe^{rt}[3 - e^m + 2me^m]}{(3 - \exp\{m\})^2},$$

where $m = P\lambda^2(e^{rt} - e^{rT})/r$. Thus $\operatorname{sgn} \partial u_i^*/\partial \lambda = \operatorname{sgn} h(m)$, where $h(m) = 3 + e^m(2m - 1)$. To see that $h(m) > 0$ for all m, let us examine $h'(m) = e^m(2m + 1)$. Now $h'(m) \gtreqless 0$ as $m \gtreqless -1/2$, so $m = -1/2$ minimizes $h(m)$. But $h(-1/2) = 3 - 2\exp\{-1/2\} > 0$. Therefore $h(m) > 0$ for all m.

Since $1/\lambda$ is the expected level of knowledge required for success, the above result implies that as the expected level of knowledge required decreases, each firm acquires knowledge at a higher rate.

The effect of increasing the discount rate r is indeterminate.

The control function $\mu_i^*(t)$ is the additional knowledge acquired by i at t. The cost of this knowledge is $(1/2)(\mu_i^*(t))^2$. To determine whether firms invest at a higher or lower rate as time progresses, we need to determine the sign of $\dot{c}_i(\mu_i^*(t))$.

PROPOSITION 2. *$c_i(\mu_i^*(t))$ is a strictly increasing and convex function of t for all $t \leqslant T$. That is, each firm's Nash equilibrium rate of investment in R & D increases (at an increasing rate) with calendar time t.*

Proof. Simply differentiate and collect terms.

COROLLARY. *We can now note that the instantaneous conditional probability of success $\lambda\mu_i^*(t)$ is an increasing function of time. That is, as time passes, it becomes increasingly likely that firm i will succeed in the next time increment dt, given that it has not yet succeeded.*

Notice that the patent value is essentially stationary while the R & D outlays are discounted. To eliminate the suspicion that this is the reason for $\dot{\mu}_i^*(t) > 0$, the same example was computed for $r = 0$, yielding the equilibrium strategies

$$u_i^*(t, z) = \frac{2P\lambda}{3 - \exp\{P\lambda^2(t - T)\}}.$$

The result remains that $\dot{\mu}_i^*(t) > 0$ for all $t \leqslant T$. This is due to the existence of the finite planning horizon. For $r > 0$, $\lim_{T \to \infty} u_i^*(t, z) = 2P\lambda e^{rt}/3$, which also increases with time. Only if the problem is autonomous and $T = \infty$ is the equilibrium strategy $u_i^*(t, z) = 2P\lambda/3$ constant over time. Lee and Wilde [8] have examined Nash equilibrium in the space of such stationary strategies. This result shows some circumstances under which stationary strategies are Nash equilibria in the larger strategy spaces $\mathscr{U}_i$.

V. THE COOPERATIVE GAME

As suggested earlier, if costs of negotiation and/or contract enforcement are not prohibitive, then the firms may choose to cooperate in the development of the innovation. By doing so, each could obtain a larger payoff (in expected value terms) than they would receive from the Nash equilibrium solution. Once the players have agreed to cooperate, they face a two-person bargaining problem to divide the total profits. Since the firms are assumed to be identical, it is only reasonable that they split the reward equally.

The strategies employed are those which maximize joint profits, $J^{12}(u_1, u_2)$. Thus far we have assumed that knowledge is privately held. But suppose that firms are able to share not only profits but research findings as well. That is, suppose knowledge is *not* embodied exclusively in the experience of the research staff but is instead transmittable and transferable through blueprints or words or formulae. Then firms 1 and 2 can form a cooperative research lab, sharing (insofar as it is feasible) the results of their individual $R \& D$ projects. Then we alter our knowledge accumulation dynamics as follows:

$$\dot{z}_i = u_i(t, z) + \gamma u_j(t, z), \tag{8}$$

where γ represents the fraction of knowledge which each firm shares with its rival. The parameter γ may represent any limitations on feasible knowledge transfer, or it may indicate the degree to which the new knowledge generated by the firms overlaps.

Now although $\mu_i(t) \in [0, B]$, the state spaces are somewhat larger, $Z_i = [0, (1 + \gamma) TB]$. The strategy space $\mathscr{U}_i$ is again the set of all bounded, continuous (in (t, z)) and Lipschitz (in z) functions defined on $[0, T] \times Z_1 \times Z_2$. These restrictions are sufficient to guarantee that any strategy pair $\in \mathscr{U}_1 \times \mathscr{U}_2$ generates a unique trajectory $(z_1(t), z_2(t))$ on $[0, T]$ through $z(0) = 0$. Thus the payoff is well defined over $\mathscr{U}_1 \times \mathscr{U}_2$.

The cooperative problem for our identical firms becomes

$$\underset{(u_1,u_2)\in\mathscr{U}_1\times\mathscr{U}_2}{\text{maximize}}\ J^{12}(u_1,u_2)=P(1-e^{-\lambda(z_1(T)+z_2(T))})$$

$$-\int_0^T e^{-rt}e^{-\lambda(z_1(t)+z_2(t))}(1/2)[(\mu_1(t))^2+(\mu_2(t))^2]\,dt$$

subject to: $\mu_i(t)=u_i(t,z(t))$, where $z(t)$ solves the system (8) with $z(0)=0$.

That is, the firms receive the patent with present value P so long as either succeeds—that is, with probability $1-e^{-\lambda(z_1(T)+z_2(T))}$.

Since firms 1 and 2 jointly choose (u_1,u_2) so as to maximize $J^{12}(u_1,u_2)$, this is just a two-state, two-control problem of dynamic optimization. We again use the dynamic programming approach.

Letting $W(t,z)$ denote the joint value function, the Bellman equation for the cooperative game is

$$W_t(t,z)+\max_{(u_1(t,z),u_2(t,z))\in[0,B]^2}[W_{z_1}(t,z)[u_1(t,z)+\gamma u_2(t,z)]$$

$$+W_{z_2}(t,z)[u_2(t,z)+\gamma u_1(t,z)]-e^{-rt}e^{-\lambda(z_1+z_2)}(1/2)[(u_1)^2+(u_2)^2]]=0. \tag{9}$$

Performing the indicated maximization yields the expression $u_i^{oo}=[W_{z_i}+\gamma W_{z_j}]\,e^{rt}e^{\lambda(z_1+z_2)}$.

Notice that the profit maximizing level of knowledge generation for player i now depends on the value of the increment to j's stock of knowledge (via the sharing arrangement) as well as upon the value of the increment to i's stock of relevant knowledge. The single Hamilton–Jacobi equation is (after simplification)

$$W_t+[(W_{z_1})^2+(W_{z_2})^2]e^{rt}e^{\lambda(z_1+z_2)}(1+\gamma^2)/2+\gamma 2W_{z_1}W_{z_2}e^{rt}e^{\lambda(z_1+z_2)}=0, \tag{10}$$

with the terminal conditions

$$W(T,z(T))=P(1-e^{-\lambda(z_1(T)+z_2(T))}). \tag{11}$$

This suggests a solution of the form $W(t,z)=b(t)e^{-\lambda(z_1+z_2)}+k(t)$, where $b(T)=-P$ and $k(T)=P$. The Hamilton–Jacobi equation becomes

$$e^{-\lambda(z_1+z_2)}[\dot b(t)+(b(t))^2\lambda^2(1+\gamma)^2e^{rt}]+\dot k(t)=0. \tag{12}$$

Since $e^{-\lambda(z_1+z_2)}$ is always positive, it must be that $k(t)\equiv P$, while $b(t)$ solves $\dot b(t)+(b(t))^2\lambda^2(1+\gamma)^2e^{rt}=0$, $b(T)=-P$. This equation is linear in $q(t)=-1/b(t)$ and thus has the unique solution

$$b(t;-P)=\frac{-P}{1-P\lambda^2(1+\gamma)^2(e^{rt}-e^{rT})/r}<0 \qquad \text{for all } t\leqslant T.$$

Thus a solution of (10) satisfying the terminal conditions (11) is

$$W(t, z; \gamma) = P - \frac{P e^{-\lambda(z_1 + z_2)}}{1 - P\lambda^2(1+\gamma)^2(e^{rt} - e^{rT})/r}.$$

Proposition 3. *Suppose $B \geqslant (1+\gamma) P\lambda e^{rT}$. Then*

$$\text{(a)} \quad u_i^{**}(t, z; \gamma) = \frac{(1+\gamma) P\lambda e^{rt}}{1 - P\lambda^2(1+\gamma)^2(e^{rt} - e^{rT})/r}, \qquad i = 1, 2,$$

is the unique cooperative equilibrium (Nash equilibrium in the one-player game) in $\mathscr{U}_1 \times \mathscr{U}_2$; and

$$\text{(b)} \quad J^{12}(u_1^{**}, u_2^{**}; \gamma) = P - \frac{P}{1 - P\lambda^2(1+\gamma)^2(1 - e^{rT})/r}, \qquad i = 1, 2.$$

The proof is methodologically identical to that of Proposition 1.

Comparative dynamics in P, T, λ, γ and r are summarized as follows:

$$\partial u_i^{**}(t, z; \gamma)/\partial P > 0 \qquad \text{for all} \quad t \in [0, T],$$

$$\partial u_i^{**}(t, z; \gamma)/\partial \lambda \begin{cases} > 0 & \text{for } t > \bar{t} \\ = 0 & \text{for } t = \bar{t} \\ < 0 & \text{for } t < \bar{t} \end{cases} \quad \text{and} \quad \partial u_i^{**}(t, z; \gamma)/\partial \gamma \begin{cases} > 0 & \text{for } t > \bar{t} \\ = 0 & \text{for } t = \bar{t} \\ < 0 & \text{for } t < \bar{t}, \end{cases}$$

where $\bar{t} = (1/r)\ln[e^{rT} - r/P\lambda^2(1+\gamma)^2]$.

Finally, $\partial u_i^{**}(t, z; \gamma)/\partial r$ is indeterminate.

Proposition 4. *The path of the equilibrium rate of investment in R & D is an increasing (and convex) function of time. The cooperative equilibrium also involves an increasing instantaneous conditional probability of success.*

Proof. Differentiate and collect terms.

Although we earlier assumed that rivals cooperated to the fullest extent possible, we can now present some justification for this assumption. Maximized expected joint profit

$$J^{12}(u_1^{**}, u_2^{**}; \gamma) = P - \frac{P}{1 - P\lambda^2(1+\gamma)^2(1 - e^{rT})/r}$$

is strictly increasing in γ so that once they decide to cooperate, the incentive is to exchange as much knowledge as possible. Sharing profits gives each firm an increased probability of receiving a payoff as well as giving the cartel an increased ability to extract profits. In addition, the ability to share research findings allows the firms to utilize the lower portions of their cost functions; that is, by splitting up the *R & D* they can generate a given level of knowledge at a lower cost than if either were to do it alone.

VI. COMPARISON OF EQUILIBRIA

Recall that noncooperative rivals each acquire new knowledge at the rate

$$u_i^*(t, z) = \frac{2P\lambda e^{rt}}{3 - \exp\{m(t)\}}$$

while cooperative rivals' equilibrium rates of knowledge acquisition are

$$\dot{z}_i^{**}(t, z; \gamma) = \frac{(1+\gamma)^2 P\lambda e^{rt}}{1 - (1+\gamma)^2 m(t)},$$

where $m(t) = P\lambda^2(e^{rt} - e^{rT})/r$. Suppose that $\gamma = 0$. This may mean that knowledge is completely embodied and nontransferable or that, for example, anti-trust laws strictly prohibit close cooperation among firms. However, firms could still cooperate to the extent of maximizing joint profits and this "tacit collusion" might be undetectable by any incompletely informed observer. The requisite wealth transfer could be made via licensing agreements.

PROPOSITION 5. *$u_i^*(t, z) > u_i^{**}(t, z; \gamma = 0)$ for all $t < T$. At $t = T$, they are equal.*

That is, Nash rivals invest in *R & D* at a uniformly higher rate than do cooperative (tacitly collusive) rivals.

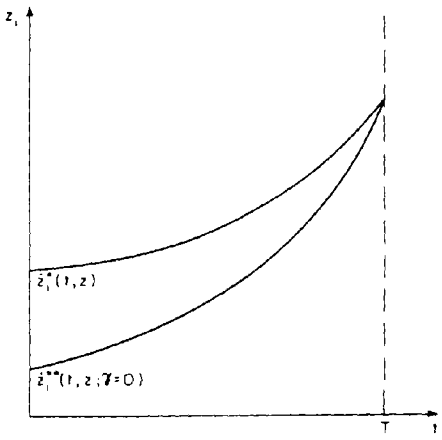


FIGURE 1

Proof. Notice that $m(t) \leqslant 0$. Now

$$u_i^*(t, z) - u_i^{**}(t, z; 0) = \frac{2P\lambda e^{rt}}{3 - \exp\{m\}} - \frac{P\lambda e^{rt}}{1 - m} > 0$$

if and only if $\exp\{m\} - 2m - 1 > 0$. Let $h(m) = e^m - 2m - 1$. Now $h(0) = 0$ and $h'(m) = e^m - 2 < 0$ for all $m \leqslant 0$. Thus as m decreases, $h(m)$ increases from zero. Thus $h(m) > 0$ for all $m < 0$ $(t < T)$. Clearly $u_i^*(T, z(T)) = P\lambda e^{rT} = u_i^{**}(T, z(T); 0)$. An illustrative graphical representation is as in Fig. 1.

Thus when the exchange of knowledge is prohibited, we can expect that an industry composed of two noncooperative rivals will complete the innovation earlier (on average) than a tacitly collusive 2-firm industry.

When firms are not prohibited from organizing a cooperative research lab and $\gamma > 0$, the results are not so clear cut. Cooperative rivals may innovate (on average) before or after noncooperative rivals, depending upon the sharing percentage γ, the expected level of required knowledge $1/\lambda$, the patent value P, terminal date T and the discount rate r. To see this, we note that the equilibrium rates of knowledge acquisition for cooperative rivals are increasing in γ. That is,

$$\partial \dot{z}_i^{**}/\partial \gamma = \partial(1 + \gamma)\, u_i^{**}(t, z; \gamma)/\partial \gamma = \frac{2P\lambda(1 + \gamma)e^{rt}}{(1 - (1 + \gamma)^2 m)^2} > 0.$$

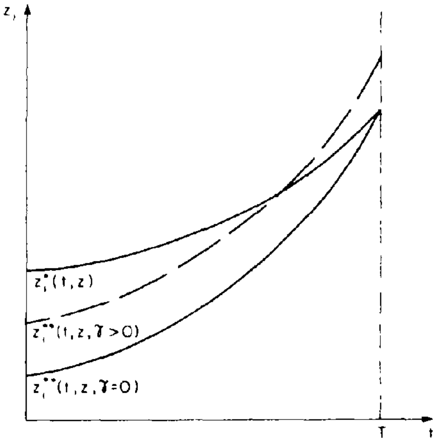


FIGURE 2

Thus for t sufficiently near T, $z_i^{**} > z_i^*$ for $\gamma > 0$. Algebraically,

$$(1 + \gamma)\, u_i^{**}(t, z; \gamma) - u_i^*(t, z) = \frac{(1 + \gamma)^2 P\lambda e^{rt}}{1 - (1 + \gamma)^2 m} - \frac{2P\lambda e^{rt}}{3 - \exp\{m\}} \gtreqless 0$$

as $h(m; \gamma) = 3(1 + \gamma)^2 - (1 + \gamma)^2 e^m - 2 + 2(1 + \gamma)^2 m \gtreqless 0$. Since $h(0; \gamma) > 0$ and $\partial h(m; \gamma)/\partial m > 0$, as m decreases, $h(m; \gamma)$ decreases. Thus there is the possibility of an intersection of the two paths, even for γ close to 1. The earliest possible intersection occurs at $\gamma = 1$. Then $2u_i^{**}(t, z; 1) - u_i^*(t, z) \gtreqless 0$ as $5 - e^m + 4m \gtreqless 0$. An intersection exists so long as $e^{rT} \geqslant 1 + 1.08r/P\lambda^2$, which covers most reasonable cases. When an intersection exists, it is not clear a priori (without specifying P, T, λ, γ, and r) whether Nash or cooperative rivals will succeed earlier on average (see Fig. 2).

VII. OPTIMALITY

The "socially optimal" level of investment in research and development is a particularly elusive notion. Several incentives exist for overinvestment in *R & D*. On an individual firm basis, the existence of a patent promises monopoly rents rather than a mere competitive return. This chance at a supernormal profit acts as a spur to innovative activity. The existence of rivals provides additional incentive to "overspend" on *R & D*, since only the first to succeed is rewarded. On an industry level, rivalry means that research effort will be duplicated as firms run parallel projects. However, it may be optimal to run parallel projects, since each additional project increases society's chances of obtaining the innovation at an earlier date.

On the other hand, the public good aspect of knowledge generation is well known. If a firm cannot expect to appropriate the full value of the knowledge it generates, its *R & D* effort will be correspondingly reduced. The patent system is an attempt to mitigate the disincentive of nonappropriability. However, the social value of an innovation may exceed the monopoly rents. Hence patents may provide only a partial remedy. If some usable technology or device is to be replaced, then the social value of the innovation may well be less than its value to the innovator. This is because the innovator wishes to redistribute another firm's wealth to itself by innovating, while society gains nothing from such redistribution.

For our particular two-firm model, the social planner's problem is identical to the two-firm cooperative model with the exception that we replace the discounted patent value P with the discounted *net* social value of the innovation Q. That is,

$$u_i^s(t, z; \gamma) = \frac{(1 + \gamma)\, Q\lambda e^{rt}}{1 - Q\lambda^2(1 + \gamma)^2(e^{rt} - e^{rT})/r}\,.$$

Notice that $u_i^s(t, z; \gamma) \gtreqless u_i^{**}(t, z; \gamma)$ as $Q \gtreqless P$. Therefore all the indeterminacies which applied to the comparison of the Nash and cooperative equilibrium strategies apply as well to the comparison of the Nash and socially optimal strategies. Nash rivals may innovate earlier or later (on average) than is socially optimal, depending upon the specific values of the parameters P, Q, T, λ, γ, and r.

VIII. KNOWLEDGE AS A PUBLIC GOOD

There are at least two ways in which a rivial firm can benefit from a particular firm's investment in *R & D*. One way is through imitation. By imitating, the rival may be able to capture some of the profits generated by an innovation. The other way involves the much-discussed public good aspect of knowledge production. In this case, the new knowledge generated by firm i is only imperfectly retained. Some of it "leaks out" to become publicly known.

In order to focus exclusively upon this form of nonappropriability, we will continue to assume that patent protection is perfect. For a two-firm game, firm i now accumulates knowledge at the rate

$$\dot{z}_i = u_i(t, z) + \rho u_j(t, z), \tag{13}$$

where ρ represents the proportion of the rival's research findings which spill over to firm i.

For simplicity, suppose that $\rho = 1$. Then $Z_i = [0, 2TB]$. With perfect patent protection and knowledge as a pure public good, firm i's payoff is

$$J^i(u_1, u_2) = P(1 - e^{-\lambda(z_1(T) + z_2(T))})/2 - \int_0^T e^{-rt} e^{-\lambda(z_1(t) + z_2(t))}(1/2)(\mu_i(t))^2\,dt.$$

The Nash equilibrium strategies must satisfy the Bellman system

$$\begin{aligned} V_t^i(t, z) + \max_{u_i(t,z) \in [0,B]} & [(V_{z_i}^i(t, z) + V_{z_j}^i(t, z))(u_i(t, z) + u_j^*(t. z)) \\ & - e^{-rt} e^{-\lambda(z_1 + z_2)}(1/2)(u_i(t, z))^2] = 0; \end{aligned} \tag{14}$$

and the terminal conditions

$$V^i(T, z(T)) = P/2 - Pe^{-\lambda(z_1(T) + z_2(T))}/2. \tag{15}$$

The candidates for an interior (for all t) Nash equilibrium may be obtained by performing the maximization indicated above. The Hamilton–Jacobi equations become

$$V_t^i + [V_{z_i}^i + V_{z_j}^i]^2 e^{rt} e^{\lambda(z_1 + z_2)}/2 + [V_{z_i}^i + V_{z_j}^i][V_{z_i}^j + V_{z_j}^j] e^{rt} e^{\lambda(z_1 + z_2)} = 0. \tag{16}$$

The terminal conditions (15) suggest a solution of the form $V^i(t, z) = b(t)e^{-\lambda(z_1+z_2)} + k(t)$, $i = 1, 2$.

Substituting the appropriate partial derivatives into Eq. (16) and simplifying implies that $k(t) \equiv P/2$ and $b(t)$ must solve

$$\dot{b}(t) + 6\lambda^2(b(t))^2 e^{rt} = 0 \tag{17}$$

with the terminal condition that $b(T) = -P/2$. Let $q(t) = -1/b(t)$. Then Eq. (17) becomes linear in $q(t)$ and has the unique solution

$$b(t; -P/2) = \frac{-P}{2 - 6P\lambda^2(e^{rt} - e^{rT})/r}.$$

Thus a solution of the system (16) satisfying the terminal conditions (15) is

$$V^i(t, z; \rho = 1) = P/2 - \frac{Pe^{-\lambda(z_1+z_2)}}{2 - 6P\lambda^2(e^{rt} - e^{rT})/r}, \qquad i = 1, 2.$$

PROPOSITION 6. *Suppose* $B \geqslant P\lambda e^{rT}$, *and suppose that knowledge is a pure public good* ($\rho = 1$). *Then*

(a) $$u_i^*(t, z; \rho = 1) = \frac{P\lambda e^{rt}}{1 - 3P\lambda^2(e^{rt} - e^{rT})/r}, \qquad i = 1, 2,$$

is the unique Nash equilibrium in $\mathscr{U}_1 \times \mathscr{U}_2$; *and*

(b) *The Nash equilibrium allocation is*

$$J^i(u_1^*, u_2^*; \rho = 1) = P/2 - \frac{P}{2 - 6P\lambda^2(1 - e^{rT})/r}, \qquad i = 1, 2.$$

The proof is methodologically identical to that of Proposition 1.

PROPOSITION 7. *Suppose that the net social value, Q, exceeds P/2, all knowledge is costlessly transferable* ($\gamma = 1$) *and knowledge is a pure public good* ($\rho = 1$). *Then each firm generates knowledge at a lower rate than is socially optimal. Consequently, innovation wll be delayed (on average) relative to the socially optimal date.*

Proof. $\dot{z}_i^*(t, z; \rho = 1) - \dot{z}_i^s(t, z; \gamma = 1) =$

$$\frac{2P\lambda e^{rt}}{1 - 3P\lambda^2(e^{rt} - e^{rT})/r} - \frac{4Q\lambda e^{rt}}{1 - 4Q\lambda^2(e^{rt} - e^{rT})/r} < 0$$

if and only if $P - 2Q + 2PQ\lambda^2(e^{rt} - e^{rT})/r < 0$. If $Q > P/2$, then since $(e^{rt} - e^{rT})/r \leqslant 0$, it follows that

$$\dot{z}_i^*(t, z; \rho = 1) - \dot{z}_i^s(t, z; \gamma = 1) < 0 \qquad \text{for all} \quad t \leqslant T.$$

That is, each firm acquires knowledge at a rate which is uniformly lower than the socially optimal rate when the social value of the innovation is at least 1/2 the value to the innovating firm.

COROLLARY. *Suppose that all knowledge is costlessly transferable* ($\gamma = 1$) *and knowledge is a pure public good* ($\rho = 1$). *Then since* $P > P/2$, *we can state that innovation by Nash rivals will occur later* (*on average*) *than the cartel date.*

Recall that for $\gamma = 0$, $\rho = 0$, Nash rivals could be expected to innovate at an *earlier* date than the cartel. Now we see that for $\gamma = 1$, $\rho = 1$, exactly the opposite is true. Therefore the determination of that market structure which most promotes innovation depends critically upon the degree to which knowledge is a public good.

It is interesting to note that if firms are able to keep their research findings private, they will not necessarily innovate earlier (on average) than if their research findings are publicly known. For although each invests at a lower rate when knowledge is a pure public good, each also benefits from the

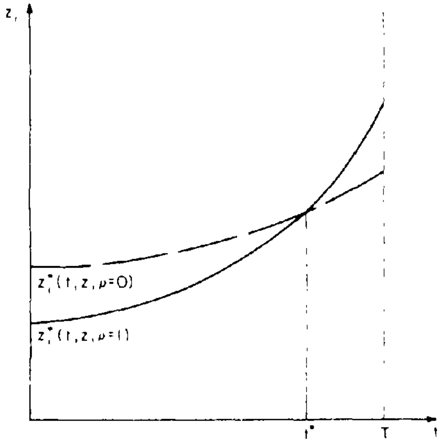


FIGURE 3

spillover of its rival's *R & D*. Thus a typical configuration is as in Fig. 3 with the intersection (if any) occurring at $t^0 = (1/r)\ln[e^{rT} - 0.455r/P\lambda^2]$.

Thus the public good aspect of knowledge generation may actually result in earlier innovation than would occur if knowledge were a private good.

IX. SUMMARY AND CONCLUSIONS

A model has been developed for rival firms engaged in industrial research and development. The firms are subject to technological uncertainty in the sense that they cannot predict with certainty the amount of *R & D* required to complete the innovation. Furthermore, each faces uncertainty regarding the date of a rival's pre-emptive strike. Several variants of rivalrous behavior are examined, including (Nash) noncooperative behavior and cooperative behavior followed by Nash bargaining over the division of the profits.

While it is clear that the Nash equilibrium and the socially optimal rates of investment in *R & D* do not coincide (neither with private nor public knowledge), it is not always clear a priori in which direction the bias acts. If knowledge is a pure public good and the social value of the innovation is high enough, then firms unambiguously underinvest as compared to the social optimum. However, if knowledge is appropriable, then although there is a misallocation of resources over time, it is not uniform. That is, with private knowledge, the paths of accumulated knowledge may cross, leaving the bias of the market (game) solution indeterminate.

In addition, we considered the question of whether a cartel can be expected to innovate earlier or later than an industry of noncooperative rivals. Again we find indeterminacies. Neither market structure dominates the other in promoting innovation. When all research findings may be privately retained, then the noncooperative rivals can be expected to innovate earlier than the cartel. But when all research findings leak out to become publicly known, this result is completely reversed.

For these reasons, the implications of this study for anti-trust policy and subsidization are not obvious. No general stance is adequate. Instead the industry environment—the degree of appropriability of knowledge, the degree of transferability of knowledge as well as the relevant parameter values—must be ascertained before a reasoned prescription can be proposed.

ACKNOWLEDGMENTS

I would like to thank Professors Morton I. Kamien, Nancy L. Schwartz and John Roberts for their comments and encouragement.

REFERENCES

1. D. BERNSTEIN, "Existence Theorems in Partial Differential Equations," Princeton Univ. Press, Princeton, N. J., 1950.
2. P. DASGUPTA AND J. STIGLITZ, Industrial structure and the nature of innovative activity, *Econ. J.* **90** (1980), 266–293.
3. P. DASGUPTA AND J. STIGLITZ, Uncertainty, industrial structure and the speed of R & D, *Bell J. Econ.* **11** (1980), 1–28.
4. A. FRIEDMAN, "Differential games," Wiley, New York, 1971.
5. M. I. KAMIEN AND N. L. SCHWARTZ, Timing of innovations under rivalry, *Econometrica* **40** (1972), 43–60.
6. M. I. KAMIEN AND N. L. SCHWARTZ, Risky R & D with rivalry, *Ann. Econ. Social Measurement* (March 1974), 267–277.
7. M. I. KAMIEN AND N. L. SCHWARTZ, On the degree of rivalry for maximum innovation activity, *Quart. J. Econ.* **90** (1976), 245–260.
8. T. LEE AND L. L. WILDE, Market structure and innovation: a reformulation, *Quart. J. Econ.* **94** (May 1980), 429–436.
9. G. C. LOURY, Market structure and innovation, *Quart. J. Econ.* **93** (1979), 395–410.
10. J. F. REINGANUM, A dynamic game of R & D: Patent protection and competitive behavior, *Econometrica*, in press..
11. J. F. REINGANUM, "Dynamic Games of R & D with Rivalry," Ph. D. Thesis, Northwestern University, August 1979.
12. L. E. RUFF, Research and technological progress in a Cournot economy, *J. Econ. Theory* (1969), 397–415.
13. F. M. SCHERER, Research and development resource allocation under rivalry, *Quart. J. Econ.* **81** (1967), 359–394.
14. K. SHELL, Inventive activity, industrial organization and economic growth, *in* "Models of Economic Growth" (J. A. Mirrlees and N. H. Stern, Eds.), Wiley, New York, 1973.
15. H. STALFORD AND G. LEITMANN, Sufficiency conditions for Nash equilibrium in n-person games, *in* "Topics in Differential Games" (Austin Blaquière, Ed.), North-Holland, New York, 1973.

JOURNAL OF ECONOMIC THEORY 25, 42–66 (1981)

Testing for Ability in a Competitive Labor Market

K. BURDETT

Economics Department, University of Wisconsin, Madison, Wisconsin 53706

AND

DALE T. MORTENSEN

Economics Department, Northwestern University,
Evanston, Illinois 40202

Received October 8, 1979; revised July 28, 1980

1. INTRODUCTION

The seminal works of Akerlof [1] and Spence [13, 14] have stimulated much research on what may be (loosely) termed the economics of information creation. Although studies in this area have considered various markets, the majority have concentrated on the creation of information about workers' abilities[1] (see, for example, Arrow [2], Riley [6], and Stiglitz [17]). The purpose of the present study is to specify conditions on the information structure that guarantee efficiency of a competitive market for worker ability information. That information structure is termed symmetric as it satisfies the following maxim: What is known to one is known to all. If the information structure is asymmetric, it does not satisfy this maxim, a competitive equilibrium will not be efficient in general. Two asymmetric information models are specified which illustrate this point. In both these models, a planner can obtain an efficient solution in principle.

There is an intimate relationship between the market for information about abilities and the labor market which utilizes this information. Surprisingly, there are very few studies in the labor economics literature that deal with a labor market where the abilities of some workers may not be known with certainty. As a labor market of this type is required to study a market for information about abilities, one is specified in some detail in Section 2. Being primarily interested in the information structure, and because we have

---

[1] Siglitz and Rothschild [8] and Wilson [18] have considered models of information creation in insurance markets.

0022-0531/81/040042-25$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

considered elsewhere a general formulation (Burdett and Mortensen [4]), only a special case of this type of labor market will be specified. The model used can be outlined briefly as follows. Suppose there are only two types of jobs available in a labor market. All workers can supply the same amount of labor services in any job of the first type (unskilled jobs). The amount of labor services a worker can supply in any job of the second type (skilled jobs) depends on that worker's ability. Although the ability of a worker is assumed to be initially unknown to all, there is a signal attached to each worker which can be costlessly observed. This signal can be used as a predictor of ability. Further information about a worker's ability can be obtained by means of a test. This test, however, is costly.[2]

Given the model outlined above, a planner faces the following allocation problem. First, the planner has to decide which workers to test. Second, the planner has to allocate workers to jobs. Tested workers can be assigned to jobs on the basis of their test scores, whereas untested workers have to be assigned to jobs on the basis of their signals. A mismatch is said to occur if a worker is allocated to a different type of job than would be the case if his or her ability were known. The greater the number tested the smaller is the number of workers mismatched; hence, the greater the expected output produced by these workers. The more workers tested, however, the greater the total cost of testing. In Section 2 conditions are specified which guarantee a planner can arrange matters to maximize the value of net output, i.e., the value of output minus the cost of testing.

In Sections 3 and 4 we investigate whether a competitive market can solve the above allocation problem. It will be shown that the assumptions made about information structure play a crucial role. In Section 3 it is assumed that the information structure is symmetric in that (a) the test score obtained by any tested worker is public information, and (b) each worker has no information above and beyond public information when deciding if the test should be taken. It will be shown that when these restrictions hold the market solution maximizes the value of net output given the information endowment.

The information structure is assumed to be asymmetric in both the two models analyzed in Section 4. In the first model considered, the test score obtained by a worker is assumed to be private information to that worker. Hence, a tested worker can masquerade as an untested worker when applying for a job. Indeed, workers who obtain "low" test scores may be motivated to claim to be untested when applying for skilled jobs. It will be shown that this dishonest behavior destroys the market for untested workers in skilled jobs, which in turn implies too many workers are tested. Thus, the value of net output is smaller than that obtained in Section 3.

---

[2] Similar labor market models have been presented previously by Rosen [9], Roy [11], and Sattinger [12]. In each of these models, however, it was assumed that information about abilities is perfect.

In the second model specified in Section 4 it is assumed that workers have special information about their own abilities before testing. Specifically, workers are assumed to be endowed with perfect information about their own abilities. Other workers and firms do not known initially the ability of a worker. In this case a worker who believes he or she has a "high" ability may be motivated to take the test to certify this fact. In such a model it can be shown that a competitive equilibrium exists and such an equilibrium can be characterized. Further, it will be demonstrated the value of net output can be increased by banning certain workers from taking the test, even though their abilities are not known.

If there is only one type of job, testing lowers the value of net output as all workers will work in the same type of job whether they are tested or not. Testing can only improve net output if there is a strictly positive probability of a mismatch, which is the case in the models considered in this study. Nevertheless, a worker who can hide his or her test result, or a worker who knows his or her ability before testing, may be motivated to take the test even though there is a zero probability this worker will be mismatched even if untested. Given either of the two asymmetric models considered, can a planner improve on the market solution? In Section 5 it is shown that a planner can achieve a first best solution in principle.

When the test score obtained by a worker is private information a planner can improve on the market solution by banning certain workers from taking the test. The planner does not need to know these workers' abilities. In this way a planner can obtain the same value of net output as generated from the market solution when the information structure is symmetric.

The planner considered in Section 2 attempts to maximize net output subject to the information endowment. Specifically, the planner knows each worker's signal and the statistical relationship between signals and abilities in the market. This planner rightly believes that workers know no more about their own abilities than the planner. In the final part of Section 5 a planner is assumed to believe rightly that each worker knows his or her own ability. It will be shown that this planner can obtain a greater value of net output than the planner considered in Section 2. Indeed, this planner, by utilizing a random testing scheme, can obtain a value of net output only marginally smaller than a planner who is endowed with perfect knowledge of each worker's ability.[3]

---

[3] Stiglitz [16] has presented a random testing scheme in a different context.

2. A Model of Labor Market Sorting

Suppose there is a large fixed number of workers in a labor force. The set of jobs at which these workers may seek employment can be partitioned into a given number of types. A worker supplies the same amount of labor services per period when employed in any job of a particular type. The specific amount a worker can supply in any job of a given type, however, depends on that worker's ability. To highlight the issues under consideration the special case will be considered where there are only two types of jobs available. Further, assume any worker can perform one (type 1) task per period in any type 1 job, whereas the number of (type 2) tasks a worker can accomplish per period in any type 2 job is termed that worker's ability. Type 1 and type 2 jobs are termed unskilled and skilled jobs, respectively.

A worker's ability is assumed to be initially unknown to all, even to that worker. Instead, attached to each worker is a signal which can be costlessly observed. For example, the signal may be based on a worker's high school record or previous work experience. The signal associated with a worker is useful to the extent it can be used to predict ability. To formalize this idea, assume each worker can be characterized by a pair of positive real numbers. Specifically, let $(a, s) \in A \times S \subset R_+^2$ denote a worker's ability/signal pair, and let the distribution of such pairs over the labor force be represented by the distribution function F. Assume A and S are compact subsets of the positive real line and suppose F has a compact support. Further, let $G(\cdot \mid s)$ denote the conditional distribution function describing the distribution of abilities for any given s.

Although the ability of a worker is initially unknown, there is a test available that reveals it. This test can be envisaged as a formal examination or a period during which a worker's performance is observed in a skilled job. The output of any firm depends on the number of type 1 and type 2 tasks performed by workers employed by that firm. Specifically, let

$$y = y(l_1, l_2) \tag{1}$$

denote the production function faced by any firm, where l_i indicates the number of type i tasks performed at the firm, $i = 1, 2$. Assume this function is increasing and strictly concave. As firms are assumed to be essentially the same, the same number of workers will be employed by each firm in equilibrium. Hence, without loss of generality, we consider one firm only and assign the number 1 to the number of workers employed by that firm.

Consider a planner endowed with knowledge of each worker's signal and the distribution of ability/signal pairs over the labor force, F. The planner's problem is to decide which workers to test and then choose the group of workers to be employed at type i jobs, $i = 1, 2$. The decision to test a worker can be based only on the information available—the worker's signal. The assignment of tested workers to jobs can be based on their known abilities, whereas only an untested worker's signal can be used when assigning him or her to a job. If a worker's signal is not a noiseless predictor of ability, an untested worker may be assigned to a job to which he or she is not best suited. Hence, if the test were costless, the planner would test all workers before allocating them to jobs. When the cost of the test is strictly positive, however, there is a significant problem determining which workers to test.

In what follows various subsets of the sets A and S will be considered. Throughout it will be assumed that any subset of $A(S)$ considered is an element of the Borel sets of $A(S)$. A solution to the planner's problem is termed a sort. Given the planner wants to employ all workers, a class of sorts can be characterized by a partition of $A \times S$ into four sets such that

(a) an untested worker with $(a, s) \in A \times S_i$ is assigned to a type i job, $i = 1, 2$;

(b) a tested worker with $(a, s) \in A_i \times S_0$ is assigned to a type i job, $i = 1, 2$;

(c) a worker with signal s is tested if and only if $s \in S_0$.

When a particular sort of the above type is used the expected number of type 1 and type 2 tasks performed at a firm can be written as

$$l_1 = l_1^u + l_1^t = \int_{A \times S_1} dF(a, s) + \int_{A_1 \times S_0} dF(a, s), \tag{2a}$$

$$l_2 = l_2^u + l_2^t = \int_{A \times S_2} a\,dF(a, s) + \int_{A_2 \times S_0} a\,dF(a, s) \tag{2b}$$

where l_i^u and l_i^t denote the number of type i tasks performed by untested and tested workers, respectively, $i = 1, 2$.

For the moment, assume the planner has decided to test only those workers with a signal in any given set S_0. Let $\Theta(S_0) \subset R_+^2$ denote the set of feasible labor inputs when the set of tested labor inputs is given by S_0, i.e., $(l_1, l_2) \in \Theta(S_0)$, if (l_1, l_2) is generated from the type of sort specified above, given S_0 is the set of tested workers, or if $(0, 0) \leqslant (l_1, l_2) \leqslant (l_1', l_2')$, where (l_1', l_2') is generated from such a sort. The following claim characterizes the set of feasible labor inputs for any given S_0.[4]

---

[4] Note that the feasible set generated assumes that each untested worker with the same signal is assigned to the same type of job. Allowing the planner to allocate untested workers with the same signal to different types of jobs does not generate any new input vectors, if F is assumed continuous. If F is not continuous, it is possible to convexify the set of feasible labor inputs by allowing the planner to place untested workers with the same signal in different types of jobs (see Burdett and Mortensen [4] for details).

Proposition 1. *If the distribution of ability/signal pairs over the labor force, F, is continuous, the set of feasible labor inputs $\Theta(S_0)$ is convex and compact for any given S_0.*

Proof. See Appendix.

Throughout the remainder of the study it will be assumed that F is continuous. To illustrate Proposition 1 assume for the moment that all workers are tested, i.e., $S_0 = S$.[5] In this case it is straightforward to show that any point on the relevant boundary of the feasible set $\Theta(S_0)$ $(=\Theta(S))$ is generated from a sort where $S_i = \varnothing$, $i = 1, 2$, $A_1 = \{a \in A \mid a < t\}$, and $A_2 = \{a \in A \mid a \geqslant t\}$, for some $t > 0$. Let $(l_1(t), l_2(t))$ denote the vector of inputs generated from such a sort. It follows from the equations of (2) that

(a) $(l_1(t), l_2(t)) \to (0, \bar{a})$ as $t \to \infty$, where $\bar{a}$ is the mean ability of all workers,

(b) $(l_1(t), l_2(t)) \to (1, 0)$ as $t \to 0$,

(c) $dl_1(t)/dl_2(t) = -t$, if $G(t \mid s) > 0$ for some s.

In Fig. 1 the relevant boundary of the feasible set is given by curve ADB.[6]

When not all the workers are tested the feasible set of labor inputs will depend on the form of the conditional distribution function of abilities for a given signal, $G(\cdot \mid s)$. Nevertheless, independent of the set of tested workers, if all workers are assigned to skilled jobs, the vector of labor inputs is given by $(0, \bar{a})$. Further, if all workers are assigned to unskilled jobs, the vector of inputs is given by $(1, 0)$. If signals are noisy predictors of abilities, the

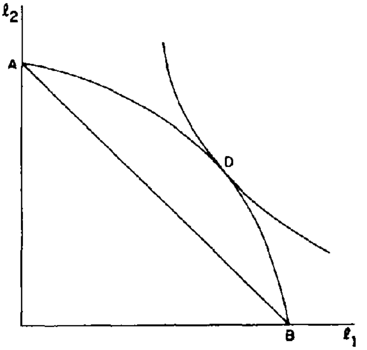


FIGURE 1

---

[5] With this restriction the model is formally equivalent to that presented by Roy [11].

[6] Given all workers are tested, point D in Figure 1 shows the input vector that maximizes output.

feasible set of labor inputs when S_0 is the set of tested workers and S_0 is a strict subset of S, is a strict subset of $\Theta(S)$. To fix ideas, consider the special case where all workers are untested and $G(a \mid s) = G(a \mid s')$ for all s and s'. i.e., $S_0 = \varnothing$ and the signal is not a predictor of ability. In this situation the slope of the relevant boundary of the feasible set is a constant equal to $\bar{a}$. This claim, which follows directly from the equations of (2), is illustrated in Fig. 1 where the relevant boundary is given by the straight line AB.

So far we have considered the situation where the set of tested workers is fixed. The number of tested workers, when S_0 is the set of tested workers, is given by

$$k = k(S_0) = \int_{A \times S_0} dF(a, s).$$

To test a worker involves the use of resources. Suppose q units of output have to be paid to test a worker. In this case $y(l_1, l_2) - qk(S_0)$ denotes net output, if $(l_1, l_2) \in \Theta(S_0)$. Suppose the planner attempts to maximize the value of net output subject to his or her information endowment, i.e., the planner knows the signals of the workers and the distribution function F.

If $k = 1$, there is only one possible set of tested workers, $S_0 = S$. Similarly, if $k = 0$, then $S_0 = \varnothing$. This is not the case when $0 < k < 1$, as many Borel subsets of S will imply k workers are tested. Let $\Xi(k) \subset R^2_+$ denote the set of feasible labor inputs when k workers are tested. It follows that $\Xi(k) = U\Theta(S_0)$, where the union is taken over all Borel sets of S which imply k workers are tested. Considering the set of feasible labor inputs for each $k \in [0, 1]$ we may define the set of sorting outcomes $\Omega \subset R^3_+$; i.e., $(l_1, l_2, k) \in \Omega$, if $(l_1, l_2) \in \Xi(k)$. Elsewhere we have demonstrated that *if the distribution of ability/signal pairs, F, is continuous, Ω is a convex subset of R^3_+* (see Burdett and Mortensen [4]). Given this result, it is straightforward to show that there exists a solution to the planner's problem. At such a solution the marginal cost of testing will equal the marginal expected gain in the value of output from testing.

Can a competitive market solve the above sorting problem? The answer to this question depends upon the assumptions made about the information structure. In the next section it will be shown that a competitive market for both types of labor and testing services will yield a solution where the value of net output is maximized, if the information structure is symmetric.

3. MARKET SORTING WITH SYMMETRIC INFORMATION

Consider the framework outlined in the previous section. Further, assume a competitive market exists in which both types of labor and testing services are traded. There are three market prices—the wage rate per type i task, w_i,

$i = 1, 2$, and the price of the test, q. There are at most three types of decision makers: firms who demand both types of labor, workers who supply labor services and demand testing services, and firms who supply the tests. Firms which demand labor services may also supply the tests. Workers are assumed to maximize their expected net incomes, whereas firms maximize their expected profits.

Two assumptions about the information structure will be maintained throughout this section.

A.1. The ability of a tested worker is known to all.

A.2. A worker's knowledge of his or her ability before testing can be described by that worker's signal s and the conditional distribution function $G(\cdot \mid s)$.

The first assumption guarantees that a tested will not masquerade as an untested worker when applying for a job. Assumption A.2 implies that a worker has no special information about his or her ability before taking the test, and thus all workers with the same signal will have the same expected return to taking the test. When these two restrictions are satisfied the information structure is termed symmetric. If at least one of them is not satisfied, the information structure is asymmetric in the sense that a worker has different information than others about his or her ability either before or after taking the test.

Without loss of generality, assume that the ability of any worker equals that worker's signal plus an unobservable error term. The special case is considered where the distribution of the error term is independent of the distribution of signals; i.e.,

$$a = s + e, \tag{3}$$

$E\{e\} = 0$, and s and e are independently distributed. Let $H(\cdot)$ and $J(\cdot)$ denote the distribution functions associated with the distribution of signals in the labor force and the error in signal term, respectively. As F was assumed to have a compact support, so do both the above distribution functions.

The ability of a tested worker is known with certainty. Hence a tested worker with ability a will receive w_1 in an unskilled job and aw_2 if employed in a skilled job. This payment system cannot be used for untested workers as their abilities are not known. This does not matter when untested workers are employed in unskilled jobs as all workers are equally productive in these jobs. Hence, assume untested workers employed in unskilled jobs are paid w_1. Suppose an untested worker employed in a skilled job is paid according to the expected ability of all workers with the same signal, i.e., sw_2 to an untested worker with signal s. After making the testing decision a worker will select to work at the type of job that yields the greatest income. For future reference, let

$$\mu(x, w) = \max(w_1, xw_2), \tag{4}$$

where $w = (w_1, w_2)$ and x equals a or s, depending on whether the worker is tested or not.

A worker with signal s will select to take the test if the expected return to taking the test, $\pi(s, w)$, is at least as great as the price of the test, q. The expected return to taking the test can be written as

$$\pi(s, w) = E\{\mu(a, w) \mid s\} - \mu(s, w). \tag{5}$$

The convexity of $\mu(a, w)$ in a and Jensen's inequality imply that $\pi(s, w)$ is nonnegative for any s and w. Further, $\pi(s, w)$ is strictly positive if and only if there is a strictly positive probability the worker with signal s would elect to work in a different type of job if his or her ability were known. The next claim characterizes the relevant properties of the expected return to testing function.

PROPOSITION 2. *The expected return to the test, $\pi(s, w)$, is a continuous function with the following properties*:

(a) $\pi(s, w) = 0$ *for any* $s < w_1/w_2$ *if and only if* $\Pr\{a > w_1/w_2\} = 0$, *and for any* $s > w_1/w_2$ *if and only if* $\Pr\{a < w_1/w_2\} = 0$.

(b) *For all* s *such that* $\pi(s, w) > 0$, $\pi(s, w)$ *is strictly increasing (decreasing) in* s *if* $s < w_1/w_2$ $(s > w_1/w_2)$.

(c) $\pi(s, w)$ *is homogeneous of degree one in* $w = (w_1, w_2)$.

Proof. Given A.1, A.2, (4), and the assumed payment structure, we have

$$\pi(s, w) = w_1 \int_0^{w_1/w_2} dG(a \mid s) + w_2 \int_{w_1/w_2}^{\infty} a\,dG(a \mid s) - \mu(s, w).$$

By virtue of (3), the appropriate change in variable yields

$$\pi(s, w) = w_1 \int_{-\infty}^{(w_1/w_2 - s)} dJ(e) + w_2 \int_{(w_1/w_2 - s)}^{\infty} (s + e)\,dJ(e) - \mu(s, w). \tag{6}$$

As $J(e)$ is continuous, $\pi(s, w)$ is differentiable except at $s = w_1/w_2$ and

$$\begin{aligned} \frac{\partial \pi(s, w)}{\partial s} &= w_2(1 - J(w_1/w_2 - s)) \geqslant 0, \qquad \text{if} \quad s < w_1/w_2 \\ &= -w_2 J(w_1/w_2 - s) \leqslant 0, \qquad \text{if} \quad s > w_1/w_2. \end{aligned} \tag{7}$$

As $\Pr(a \leqslant w_1/w_2 \mid s) = J(w_1/w_2 - s)$, the assertions made follow from inspection of (7).

The economics underlying the demand for testing services can be summarized as follows. A worker with signal $s < w_1/w_2$ has an incentive to take the test only to the extent that there is a strictly positive probability his or her ability will turn out to be greater than w_1/w_2. Only if this worker's ability is revealed to be greater than w_1/w_2 will more be earned as a tested worker. Further, the assumption that the distribution of the error in signal term is independent of the worker's signal implies the probability a worker with signal $s < w_1/w_2$ will turn out to have an ability greater than w_1/w_2 increases as s increases. A worker with signal $s > w_1/w_2$ may be willing to pay for the test as the minimum wage paid to tested workers is w_1, and this will occur only if the worker's ability is no greater than w_1/w_2. A worker with a signal so great that there is zero probability ability will be less than w_1/w_2 will not choose to take the test if $q > 0$.

The shape of the expected return to testing function specified in Proposition 2 is to a large extent determined by the assumption that the distribution of error in signal terms is independent of the distribution of signals. If these distributions are not independent, different expected return to testing functions can be generated, depending on the specific assumption made. Nevertheless, the logic underlying the shape of the function will remain the same. Specifically, expected return is positive at a given signal if and only if there is a risk of a mismatch when untested and peaks when the risk of mismatch is maximum.

A worker with signal s will choose to take the test if and only if $\pi(s, w) \geqslant q$. Utilizing Proposition 2 it follows that all workers with a signal s such that $s_1 \leqslant s \leqslant s_2$ will take the test, where $\pi(s_i, w) = q$, $i = 1, 2$. This is illustrated in Fig. 2. Any worker with a signal not in the interval $[s_1, s_2]$ will select not to take the test. Note that this result implies that, except for a set of measure zero, all workers with a given signal will take the test, or all

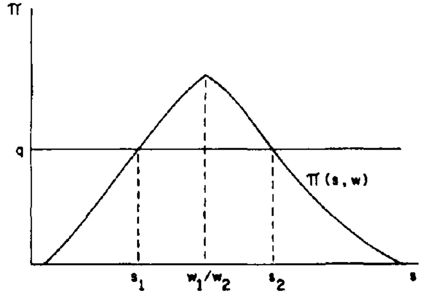


FIGURE 2

workers with this signal will not take the test. Thus, the expected ability of untested workers with signal s is s. As the number of workers is assumed to be infinite, the law of large numbers guarantees that the difference between the expected and actual number of tesks accomplished by untested workers with signal s can be ignored. Hence, the assumed payment structure to untested workers implies that untested workers are paid according to their expected ability in both types of jobs.

Suppose w_1, w_2, and q are the prices called in the market. Workers, attempting to maximize their net incomes, will sort themselves into tested and untested workers, and workers employed in skilled and unskilled jobs. The next Proposition summarizes the results shown above and thus no proof will be presented.

PROPOSITION 3. *Given the vector of wage rates w and the price of the test q, workers maximize their own net incomes if*

$$\begin{aligned}
S_0 &= S_0(w, q) = \{s \in S \mid \pi(s, w) \geqslant q\},\\
S_1 &= S_1(w, q) = \{s \in S \mid s \notin S_0 \text{ and } w_1 > sw_2\},\\
S_2 &= S_2(w, q) = \{s \in S \mid s \notin S_0 \text{ and } w_1 \leqslant sw_2\},\\
A_1 &= A_1(w, q) = \{a \in A \mid w_1 > aw_2\},\\
A_2 &= A_2(w, q) = \{a \in A \mid w_1 \leqslant aw_2\}.
\end{aligned}$$

By virtue of Propositions 2 and 3, the demand for the test and the supplies of both types of labor can be expressed as follows:

$$k^d(w, q) = \int_{A \times S_0} dF(a, s) = \int_{s_1}^{s_2} dH(s), \tag{8}$$

$$\begin{aligned}
l_1^s(w, q) &= \int_{A \times S_1(w,q)} dF(a, s) + \int_{A_1(w,q) \times S_0(w,q)} dF(a, s) \qquad (9a)\\
&= \int_{s_1}^{s_2} \int_0^{w_1/w_2} dG(a \mid s)\, dH(s) + \int_0^{s_1} dH(s),
\end{aligned}$$

and

$$\begin{aligned}
l_2^s(w, q) &= \int_{A \times S_2(w,q)} a\,dF(a, s) + \int_{A_2(w,q) \times S_0(w,q)} a\,dF(a, s)\\
&= \int_{s_1}^{s_2} \int_{w_1/w_2}^{\infty} a\, dG(a \mid s)\, dH(s) + \int_{s_2}^{\infty} s\, dH(s).
\end{aligned}$$

The properties of these functions are summarized below.

PROPOSITION 4. *The proportion of the workers who demand the test,* $k^d(w, q)$, *is*

(a) *homogeneous of degree zero and continuous;*

(b) *a decreasing function of the price of the test;*

(c) *an increasing function of the wage rates, if their ratio is held constant;*

(d) *an increasing function of the uncertainty of the signal as a predictor of ability.*

Proof. $\pi(s, w)$ is homogeneous of degree one in w for fixed s. Hence, the roots of $\pi(s, w) = q$, s_1 and s_2, are invariant with respect to scalar multiplication of all prices. This fact and (8) imply (a). Claim (b) is implied by Proposition 2 and (8). Claim (c) follows from claim (a) and claim (b). Let $J_1(\cdot)$ and $J_2(\cdot)$ denote the distribution functions associated with two random variables which have the same expected value. Rothschild and Stiglitz [7] define the random variable associated with $J_1(\cdot)$ to be riskier than the one associated with $J_2(\cdot)$, if

$$T(y) = \int_{-\infty}^{y} |J_1(e) - J_2(e)|\, de \geqslant 0, \qquad \text{for all } y,$$

and $\lim_{y\to\infty} T(y) = 0$. Integration of (6) by parts yields

$$\pi(s, w) = sw_2 - \mu(s, w) + w_2 \int_{-\infty}^{(w_1/w_2 - s)} J(e)\, de.$$

Let $\pi_i(s, w)$ equal the right hand side when $J_i(\cdot)$ replaces $J(\cdot)$, $i = 1, 2$. It follows immediately that $\pi_1(s, w) \geqslant \pi_2(s, w)$, if the random variable associated with $J_1(\cdot)$ is riskier than the other. This establishes claim (d).

PROPOSITION 5. *The labor supply functions,* $l_1^s(w, q)$ *and* $l_2^s(w, q)$, *are nonnegative, continuous, and homogeneous of degree zero. Further,* $l_1^s((0, w_2), q) = l_2^s((w_1, 0), q) = 0$.

Proof. The assertions made are implied by the equations of (9) as the roots of $\pi(s, w) = q$ are continuous and homogeneous of degree zero.

To specify fully the market, the demand for labor functions and the supply of testing services need to be derived. As the derivation of these functions is straightforward, only a brief outline will be presented. Let $l_1^d(p, w)$ and $l_2^d(p, w)$ denote the demand for type 1 and type 2 tasks, respectively, where p indicates the price per unit of the output of the firm. The restrictions placed on the production function imply both these demand functions are *continuous, decreasing, and homogeneous of degree zero. Further, if* $w_1 > 0$, *then* $l_1^d(p, (w_1, 0)) = 0$, *and* $w_2 > 0$ *implies* $l_2^d(p, (0, w_2)) = 0$.

Suppose the suppliers of the tests are price takers who utilize convex technologies. Let $k^s(q)$ denote the number of workers testing firms are willing to test when the price is q. It follows that $k^s(q)$ is an increasing function such that $k^s(0) = 0$. The lack of externalities and the competitive nature of the testing market imply q can be interpreted as both the private and social marginal cost of testing when $k^s(q)$ have been tested.

Utilizing the above and Propositions 4 and 5 it is straightforward to establish the existence and efficiency of a competitive equilibrium. As the nature and form of the proof of this claim are well known to economists, none will be presented.

PROPOSITION 6. *For any given* $p > 0$, *there exist three prices* w_1^0, w_2^0, *and* q^0 *that equate the supplies and demands for both types of labor services and the supply and demand for testing services. Further, at the competitive equilibrium the value of net output is maximized, given the information endowments.*

4. INCENTIVES AND ASYMMETRIC INFORMATION

The existence and efficiency of the market solution to the sorting problem, demonstrated in the previous section, crucially depend on the restrictions A.1 and A.2. When either of these restrictions is replaced the information structure becomes asymmetric in the sense that a worker has different information than others about his or her ability, either before or after taking the test. In this section two models are presented. In the first model A.1 is replaced, whereas in the second model A.2 is changed. It will be shown that radically different problems are encountered when considering models with an asymmetric information structure than those faced in the previous section.

4(a) *Asymmetric Information Model 1*

Assumption A.1 implies the test score obtained by any worker is public information. At first blush, this restriction appears to violate the notion that exclusive property rights are necessary for the efficiency of a competitive equilibrium. On second thought, this paradox can be resolved by distinguishing between the test itself and the information it yields. The test itself is a private good produced by scarce resources. The information it yields, however, is a public good which is free to the extent it can be costlessly transmitted. Hence, efficiency is obtained when there is free access to workers' test scores.

In this section it is shown that free access to test results is generally a necessary condition for efficiency of the market solution. Although a competitive equilibrium will be shown to exist when test scores are private information, the equilibrium has the property that taking the test is a requirement for employment in a skilled job. This leads to too many workers being tested at the market solution.

Consider the model outlined in the previous section with one important exception. Assume A.1 is replaced by the following assumption.

A.1′. The ability of a tested worker is private information to that worker.

A worker, in this case, can falsely claim to be untested when applying for a job. If a worker claims to be tested, he or she must provide proof of ability. This restriction rules out the possibility of either an untested worker falsely claiming to be tested, or a tested worker incorrectly reporting ability.

The nature of the adverse selection problem that emerges when A.1′ holds is easily illustrated. Consider the options open to a tested worker when the prices called are consistent with an equilibrium when A.1 and A.2 hold. Assume this worker has signal s and obtained a score of a in the test. If this worker reveals that he or she has been tested, w_1 and aw_2 will be earned in an unskilled and skilled job, respectively. If this worker claims to be untested, w_1 and sw_2 will be earned in an unskilled and skilled job respectively. It follows that this worker will strictly prefer to reveal the test has been taken if $a > s$ and $a > w_1/w_2$. Suppose, for example, the worker's signal is such that $sw_2 > w_1$, and thus the worker will prefer to work in a skilled job when claiming to be untested. Suppose this worker takes the test and has a revealed ability equal to $a < s$. This worker will obtain a greater income by claiming to be untested. Hence, paying $w_2 s$ to workers in skilled jobs who have signal s and claim to be untested is generally inconsistent with a competitive equilibrium. The payment to this group of workers is based on the belief that their average ability is equal to their common signal s, whereas their expected ability is $E(a \mid s \text{ and } a < s)$. Equivalently, the signal is biased upward as a predictor of ability among this group of workers.

Can the signal be appropriately discounted to correct for this self selection? To answer this question consider the following wage structure: Pay w_1 in an unskilled job and aw_2 in a skilled job to workers with ability a who claim to be tested, and pay w_1 in an unskilled job and $\phi(s)w_2$ in a skilled job to workers with signal s who claim to be untested, where $0 \leqslant \phi(s) \leqslant s$. It is straightforward to check that no other type of nondiscriminatory payment system could possibly support an equilibrium.

Let $\rho(\phi(s), w)$ denote the expected return to testing in the situation envisaged to a worker with signal s. Modifying (6) to satisfy A.1′ yields

$$\begin{aligned}\rho(\phi(s), w) &= w_1 \int_0^{w_1/w_2} dG(a \mid s) + w_2 \int_{w_1/w_2}^{\infty} a\, dG(a \mid s) - w_1, \quad \text{if} \quad w_1 > w_2\phi(s)\\ &= \phi(s)\, w_2 \int_0^{\phi(s)} dG(a \mid s) + w_2 \int_{\phi(s)}^{\infty} a\, dG(a \mid s) - w_2\phi(s),\\ &\qquad \text{if} \quad w_1 \leqslant w_2\phi(s).\end{aligned} \tag{10}$$

Next we demonstrate the relationship between $\rho(\phi(s), w)$ and $\pi(s, w)$ defined in (6).

Proposition 7. *$\rho(\phi(s), w) \geqslant q$ for any $s \geqslant s_1$, where s_1 is the smaller root of $\pi(s, w) - q = 0$.*

Proof. Using (3) to change the variable of integration in (10) implies

$$\begin{aligned}\rho(\phi(s), w) &= w_1 \int_{-\infty}^{(w_1/w_2 - s)} dj(e) + w_2 \int_{(w_1/w_2 - s)}^{\infty} (s+e)\, dJ(e) - w_1,\\ &\qquad \text{if} \quad w_1 > w_2\phi(s)\\ &= w_2\phi(s) \int_{-\infty}^{\phi(s)-s} dJ(e) + w_2 \int_{\phi(s)-s}^{\infty} (s+e)\, dJ(e) - w_2\phi(s),\\ &\qquad \text{if} \quad w_1 \leqslant w_2\phi(s).\end{aligned}$$

From inspection of (6) it follows that $\rho(\phi(s), w) = \pi(s, w)$, if s is such that $w_1 \geqslant s w_2$, and $\rho(\phi(s), w) \geqslant \pi(w_1/w_2, w)$ otherwise. This is illustrated in Fig. 3. Note that $\rho(s, w)$ is a constant if $w_1 \leqslant s w_2$, and $\rho(s, w) \leqslant \rho(\phi(s), w)$ for any $\phi(\cdot)$ such that $0 \leqslant \phi(s) \leqslant s$. This establishes the claim.

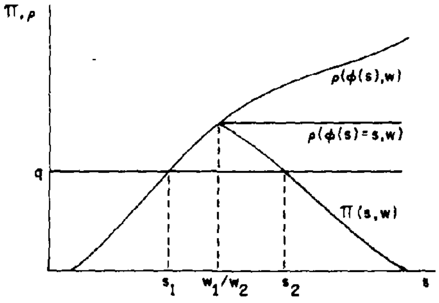


Figure 3

The above Proposition implies that if a worker with a given signal selects to take the test all workers with a greater signal will also take the test. A payment function $\phi(\cdot)$ is compatible with a market equilibrium if either (a) the expected ability of workers with signal s, who claim to be untested and work in skilled jobs, is $\phi(s)$ for any s; or (b) no workers who claim to be untested seek employment in a skilled job. A tested worker with ability/signal pair (a, s) will strictly prefer to claim to be untested when applying for a job if and only if $w_2\phi(s) > \max(w_1, aw_2)$; or equivalently, $\phi(s) > \max(w_1/w_2, a)$. Thus the ability of such a worker must be less than $\phi(s)$, and therefore the expected ability all such workers with the same signal will be less than $\phi(s)$. This implies the only payment function compatible with an equilibrium must rule out workers who claim to be untested applying for skilled jobs. It is straigtforward to show that if the payment function $\phi(\cdot)$ is such that $\phi(s) \leqslant w_1/w_2$ for all s, no worker will prefer to seek employment in a skilled job on the basis of his or her signal. It also implies no tested worker will masquerade as an untested worker.

Utilizing the above it follows that the actions of the workers at a market equilibrium will induce the following sort when A.1′ and A.2 hold:

$$\begin{aligned}
S_0 &= \{s \in S \mid \rho(\phi(s), w) \geqslant q\} = \{s \in S \mid s \geqslant s_1\},\\
S_1 &= \{s \in S \mid \rho(\phi(s), w) < q \text{ and } w_1 \geqslant w_2\phi(s)\}\\
&= \{s \in S \mid s < s_1 \text{ and } w_1 \geqslant w_2\phi(s)\},\\
S_2 &= \{s \in S \mid \rho(\phi(s), w) < q \text{ and } w_1 < w_2\phi(s)\} = \varnothing,\\
A_1 &= \{a \in A \mid w_1 \geqslant aw_2\},\\
A_2 &= \{a \in A \mid w_1 < aw_2\},
\end{aligned}$$

where $\phi(s) \leqslant w_1/w_2$ and s_1 is the smallest root of $\pi(s, w) - q = 0$.

Utilizing the above we may specify the supplies of both types of labor services and the demand for testing as follows:

$$\begin{aligned}
\bar{l}_1^s(w, q) &= \int_0^{s_1}\int_0^{\infty} dG(a \mid s)\, dH(s) + \int_{s_1}^{\infty}\int_0^{w_1/w_2} dG(a \mid s)\, dH(s),\\
\bar{l}_2^s(w, q) &= \int_{s_1}^{\infty}\int_{w_1/w_2}^{\infty} a\, dG(a \mid s)\, dH(s),\\
\bar{k}^d(w, q) &= \int_{s_1}^{\infty} dH(s).
\end{aligned}$$

By virtue of Proposition 4, the above equations are continuous and homogeneous of degree zero. Further, $\bar{l}_1^s((0, w_2), q) = \bar{l}_2^s((w_1, 0), q) = 0$. These facts, and the specification of the demands for both types of labor services and the supply of tests made in Section 3, imply it is straightforward to establish that a competitive equilibrium exists. At an equilibrium only tested workers will be employed in skilled jobs. The market clearing prices imply that no tested worker will seek employment in a skilled job masquerading as an untested worker.

Too many workers are tested in equilibrium when assumption A.1′ and A.2 obtain; i.e., the private benefit to taking the test exceeds the social benefit for some workers and is no greater for the others. To demonstrate this claim let (w_1^I, w_2^I) denote the equilibrium wage rate vector and q^I the price of the test when A.1′ and A.2 hold. From the results presented above it follows that at the equilibrium a worker will select to be tested if and only if his or her signal s is such that $s \geqslant s^I$, where $\rho(\phi(s^I), (w_1^I, w_2^I)) = q^I$. Hence, all workers with a signal s such that $s \geqslant s^I$ and $\Pr(a \leqslant w_1^I/w_2^I \mid s) = 0$ will take the test even though there is zero probability such workers will be mismatched. If all such workers are banned from taking the test the value of net output is increased. From (11) it follows that the private benefit to taking the test is at least as great as the social benefit for all workers.

4(b) *Asymmetric Information Model 2*

So far it has been assumed that a worker has no special information, above and beyond public information, about his or her own ability. It appears reasonable to argue, however, that workers often do have special information about their own abilities. For example, a worker may rightly believe that his ability is greater than the average over all workers with the same signal. This may lead this worker to certify his ability by taking the test, whereas another worker with the same signal may not select to take the test. In this subsection the special case is considered where all workers know their own abilities before testing. Formally, we consider the model outlined in Section 3 when A.2 is replaced by the following assumption:

A.2′. Each worker is endowed with perfect knowledge of his or her own ability.

Although each worker is assumed to know his or her own ability before testing, firms and other workers do not have this information. Further, there are several reasons to suspect that a worker will not truthfully reveal his or her ability to a prospective employer, especially if the wage paid is a positive function of ability. In this case some workers may be willing to certify their abilities by taking the test. Specifically, a worker will choose to take the test if income as a tested worker, minus the cost of the test, is at least as great as income as an untested worker. Note that in this case a worker is not taking a gamble by taking the test as the worker knows beforehand the score to be obtained.

Suppose, for the moment, that the wage payment system is the same as that used in Section 3; i.e., a worker with ability/signal pair (a, s) receives w_1 in an unskilled job, sw_2 as an untested worker in a skilled job, and aw_2 as a tested worker in a skilled job. This worker will prefer to take the test if and only if $aw_2 \geqslant \max(w_1, sw_2) + q$. Assuming the worker's signal is such that $sw_2 > w_1$, it follows that the expected ability of all workers with the same signal who select not to take the test is $E\{a \mid s \text{ and } a < s + q/w_2\}$. This implies the signal of an untested worker in a skilled job is an (upward) biased predictor of ability, and thus the payment system assumed above cannot support a competitive equilibrium for reasons stated in Section 4(a).

In an attempt to correct for this bias suppose firms utilize the following payment system. An untested worker with signal s is paid w_1 in an unskilled job and $w_2(s - \lambda)$ in a skilled job, whereas a tested worker with ability a recieves w_1 in an unskilled job and aw_2 in a skilled job. Assumption A.2′ implies a worker with ability/signal pair (a, s) will select to take the test, and thus certify ability, if and only if

$$\begin{aligned} a \geqslant \quad & w_1/w_2 + q/w_2, \text{ if } \quad s \leqslant w_1/w_2 + \lambda, \\ \geqslant \quad & s - \lambda + q/w_2, \text{ if } \quad s > w_1/w_2 + \lambda. \end{aligned} \tag{11}$$

By virtue of (3) and (11), the expected ability of all untested workers with signal s can be written as

$$E\{a \mid s \text{ and } a < s - \lambda + q/w_2\} = s + E\{e \mid e < q/w_2 - \lambda\}. \tag{12}$$

Hence, untested workers with signal s who work in skilled jobs will be paid according to the expected ability of the group if and only if

$$z(q/w_2, \lambda) = E\{e \mid e < q/w_2 - \lambda\} - \lambda = 0. \tag{13}$$

We are now in a position to establish that there exists a $\lambda \geqslant 0$ which implies untested workers in skilled jobs will be paid according to their expected abilities. Before this claim is made consider the following definitions. Let $\bar{e} = \sup_{J(e) < 1} e$, and $\mathbf{e} = -\inf_{J(e) > 0} e$. As $E\{e\} = 0$, both $\bar{e}$ and $\mathbf{e}$ are both strictly positive when $J(\cdot)$ is a nondegenerate distribution function.

PROPOSITION 8. *For fixed* $q/w_2 > 0$, *there exists a* $\lambda^* = \lambda(q/w_2)$ *such that* $z(q/w_2, \lambda^*) = 0$, *and* $0 \leqslant \lambda^* \leqslant \mathbf{e}$.

Proof. From (13) it follows that $z(q/w_2, 0) = 0$, if $q/w_2 \geqslant \bar{e}$, and $z(q/w_2, 0) < 0$, if $0 < q/w_2 < \bar{e}$. Further, (13) also implies that $z(q/w_2, \mathbf{e}) \geqslant 0$ for any $q/w_2 > 0$. This establishes the claim as $z(q/w_2, \lambda)$ is continuous in λ if $0 \leqslant \lambda \leqslant \mathbf{e}$.

Note that Proposition 8 specifies the form of the payment system that is compatible with a competitive equilibrium if the market clearing prices can be found. A difficulty in establishing these market clearing prices is encountered due to the fact that for fixed $q/w_2 > 0$ more than one λ may imply $z(q/w_2, \lambda) = 0$, i.e., λ^* is not unique. The uniqueness of λ^* for any fixed $q/w_2 > 0$ can be guaranteed if the "right" assumption is made about the form of the distribution function $J(\cdot)$.[7] Specificaly, if the assumptions made about the form of $J(\cdot)$ imply that $z(q/w_2, \lambda)$ is strictly increasing in λ, then λ^* can be guaranteed to be unique. In what follows it will be assumed that the form of $J(\cdot)$ is such that λ^* is unique for any fixed $q/w_2 > 0$. Although it is relatively straightforward to establish an equilibrium exists when λ^* is not unique, it will not be attempted here as it would be a distraction from the major issues under consideration. From (13) it follows that if λ^* is unique, $\lambda(q/w_2)$ is a nondecreasing function.

Let $k(w, q, s)$ denote the proportion of workers with signal s who select to be tested when w is the vector of wage rates, and q the price of the test. Using (3), (11), and (12) it follows that

$$
\begin{aligned}
k(w, q, s) &= \int_{(w_1/w_2 + q/w_2 - s)}^{\infty} dJ(e), && \text{if} \quad s < s^*(w, q)\\
&= \int_{(q/w_2 - \lambda(q/w_2))}^{\infty} dJ(e) = k(w, q, s^*(w, q)), && \text{if} \quad s \geqslant s^*(w, q),
\end{aligned}
\tag{14}
$$

where

$$
s^*(w, q) = w_1/w_2 + \lambda(q/w_2). \tag{15}
$$

Let $\mathbf{s}$ be the largest signal such that $\Pr(a \geqslant w_1/w_2 + q/w_2 - \mathbf{s} \mid \mathbf{s}) = 0$. All workers with a signal no greater than $\mathbf{s}$ will choose not to certify their abilities; i.e., $k(w, q, s) = 0$, if $s \leqslant \mathbf{s}$. When $\mathbf{s} < s < s^*(w, q)$, $k(w, q, s)$ is an increasing function of s. Further, if $s \geqslant s^*(w, q)$, the same proportion of all workers with the same signal will choose to take the test.

The properties of (14) can be written as

$$
\partial k(w, q, s)/\partial w_1 < 0, \quad \partial k(w, q, s)/\partial w_2 > 0, \quad \text{and} \quad \partial k(w, q, s)/\partial q < 0, \tag{16}
$$

if $s < \mathbf{s}$. Similarly, using (15) it can be shown that

$$
\partial s^*/\partial w_1 > 0, \quad \partial s^*/\partial w_2 \gtrless 0 \text{ as } w_1 + \lambda'(q/w_2)q \lessgtr 0, \text{ and } \partial s^*/\partial q < 0. \tag{17}
$$

The interpretations of the results presented in (16) and (17) are obvious and will not be stated. When A.1 and A.2′ hold, the supply of both types of labor services and the demand for the test can be written as:

---

[7] It is straightforward to check that if the second derivative of $E\{e \mid e \leqslant t\}$ with respect to t has the same sign for any t, then λ^* is unique.

$$\bar{l}_1^s(w, q) = \int_0^{(w_1/w_2 + \lambda(q/w_2))} (1 - k(w, q, s))\, dH(s),$$

$$\bar{l}_2^s(w, q) = \int_0^{(w_1/w_2 + \lambda(q/w_2))} \int_{(w_1/w_2 + q/w_2 - s)}^{\infty} (s + e)\, dJ(e)\, dH(s)$$

$$+ \int_{(w_1/w_2 + \lambda(q/w_2))}^{\infty} dH(s), \tag{18}$$

$$\bar{k}^d(w, q) = \int_0^{(w_1/w_2 + \lambda(q/w_2))} k(w, q, s)\, dH(s)$$

$$+ k(w, q, s) \int_{(w_1/w_2 + \lambda(q/w_2))} dH(s). \tag{19}$$

Note that the equations of (18) and (19) are continuous, homogeneous of degree zero, and $\bar{l}_1^s((0, w_2), q) = \bar{l}_2^s((w_1, 0), q) = 0$.

The demand for both types of labor services and the supply of tests are the same as specified in Section 2. Using the properties of (18) and (19), and the properties of the demand for labor services and the supply of tests previously specified, the existence of a competitive equilibrium can be readily established. At such an equilibrium all workers with a signal $s \geqslant s^*((w_1^{II}, w_2^{II}), q^{II})$ will work in skilled jobs, where w_1^{II}, w_2^{II}, and q^{II} are market clearing prices. Considering those workers with a signal $s < s^*((w_1^{II}, w_2^{II}), q^{II})$, only those workers with ability at least as great as $w_1^{II}/w_2^{II} + q^{II}/w_2^{II}$ will work in skilled jobs; the others will seek employment in unskilled jobs.

5. WELFARE IMPLICATIONS

In Section 3 it was demonstrated that when the information structure is symmetric the competitive equilibrium generates a sort that maximizes the value of net output, given the information endowment. Thus, a planner who is endowed with knowledge of each worker's signal and the distribution of ability/signal pairs, F, cannot improve on the market solution when A.1 and A.2 hold. In this section it will be shown that a planner can improve on the market solution when the information structure is asymmetric.

Suppose assumptions A.1 and A.2′ hold, i.e., tested workers can masquerade as untested workers. In Section 4(a) it was shown that in this case the market solution generates a sort of the type described in Section 2.

The sort generated, however, is different from the one that maximizes the value of net output when the information structure is symmetric. Specifically, it was shown that too many workers are tested when restrictions A.1′ and A.2 are used; i.e., the private benefit exceeds the public benefit for some workers and is no smaller for the other workers. The externality present when test scores are not public information is similar to that identified by Akerlof [1] in his analysis of the market for "lemons." Adverse selection by the less able workers with a given signal destroys the market for untested workers in skilled jobs just as adverse selection by the owners of "lemons" destroyed the market in Akerlof's model.

Can a planner arrange matters to improve on the market outcome when A.1′ and A.2 hold? In what follows it will be shown that a planner can induce a sort that maximizes the value of net output when the information structure is symmetric. Thus, a planner can improve on the market solution when assumptions A.1′ and A.2 hold. Let (w_1^0, w_2^0, q^0) denote the vector of equilibrium prices when A.1 and A.2 hold. From the results presented in Section 3 it follows that at this equilibrium a worker with signal s will take the test if and only if $s \in [s_1, s_2]$, where $\pi(s_i, (w_1^0, w_2^0)) = q^0$, $i = 1, 2$. If the same prices are called when A.1′ and A.2 are used a worker with signal s will take the test if and only if $s \geqslant s_1$. Suppose a planner calls these prices with one exception. Assume the planner bans all workers with a signal greater than s_2 from taking the test. It is straightforward to check that if A.1′ and A.2 hold, this strategy by the planner induces a sort which is identical to that generated by a competitive equilibrium when the information structure is symmetric. Note that the planner cannot improve on the value of net output generated from this strategy. A simpler and more practical way for the planner to correct matters when A.1′ and A.2 hold is to centralize test score information. If firms can costlessly check if a worker has been tested or not, the situation is formally identical to the symmetric information case. Of course, this action by a planner changes assumption A.1′ back to A.1.

In Section 2 the planner was endowed with knowledge of each worker's signal and the distribution of ability/signal pairs in the market, F. This planner correctly believed that workers had no more than public information about their own abilities before testing. Suppose a second planner is endowed with the same information with the exception that this planner rightly believes that workers under his control know their own abilities. Note that the planner does not know any worker's ability before testing; only that workers know their own abilities. Hence, the second planner is endowed with more information than the planner considered in Section 2. Can the planner utilize this extra information to obtain a greater value of net output than that obtained by the planner in Section 2? Below it will be shown this question can be answered in the affirmative.

Let (w_1^*, w_2^*) denote the vector of competitive equilibrium wage rates when the ability of each worker is known, i.e., the set of tested workers is $S_0 = S$. Suppose a planner is endowed with knowledge of each worker's signal and the distribution function F. This planner is aware that each worker knows his or her own ability. Assume this planner call the wage vector (w_1^*, w_2^*) and tells each worker to report his or her ability when applying for a job. Hence, a worker who claims to have ability $\tilde{a}$ will seek employment in a skilled job if and only if $w_1^* \leqslant \tilde{a} w_2^*$, otherwise this worker will work in an unskilled job. In an attempt to keep workers honest the planner announces that δ percent of workers will be tested at a cost of q to each worker tested. The workers chosen to be tested will be selected at random. Any worker who is found to have falsely reported ability will be forced to pay a penalty, b, and then be paid according to his or her true ability.

Given the planner's strategy outlined above, consider the problem faced by a worker with characteristics (a, s). The expected income of this worker is a function of the ability reported to the planner. Specifically,

$$\begin{aligned}\psi(\tilde{a}, a, s) &= \max(w^*, \tilde{a} w^*) - \delta q, \qquad \text{if} \quad \tilde{a} \leqslant a \\ &= (1 - \delta) \max(w^*, \tilde{a} w^*) + (\max(w^*, a w^*) - q - b), \qquad \text{if} \quad \tilde{a} > a,\end{aligned} \tag{20}$$

denote the expected income to a worker with characteristics (a, s) who claims to have ability $\tilde{a}$. Note that this worker cannot claim to have an ability greater than $s + \bar{e}$ as there is zero probability this could be true, i.e., $\tilde{a} \leqslant s + \bar{e}$, where $\bar{e}$ is defined in Section 4(b). A worker with characteristics (a, s) prefers to reveal truthfully his or her ability if and only if

$$\delta b \geqslant (1 - \delta)[(s + \bar{e}) w_2^* - \max(w_1^*, a w_2^*)]. \tag{21}$$

Since the right-hand side of (21) is no greater than $w_2^*(\bar{e} - e)$, it follows that for any δ there exists an $h(\delta)$ such that

$$\delta h(\delta) = (1 - \delta)[(s + \bar{e}) w_2^* - \max(w_1^*, a w_2^*)]$$

for all (a, s). It is straightforward to check that $h(\delta)$ is an increasing function for $0 < \delta < 1$. Hence, if the planner can inflict any penalty on the workers, it is possible to force all workers to reveal truthfully their abilities by randomly testing δ percent of them, for any δ such that $0 < \delta < 1$. Of course, the planner may be restricted in the penalty that can be inflicted. This will create a lower bound on the percent the planner can test to generate truthful behavior.

APPENDIX

Proof of Proposition 1. Let $\gamma(A)$ denote the Borel σ-algebra of the set A. From the equations of (2) it follows that, for any fixed S_0, the set $\Theta(S_0)$ is constructed from the sum of two sets: the set of feasible tested labor inputs and the set of feasible untested labor inputs. Hence, if both these sets are convex, $\Theta(S_0)$ is convex. Given any fixed S_0, (l'_1, l'_2) is an element of the set of feasible tested labor inputs if there exists two disjoint sets $A_1, A_2 \in \gamma(A)$ such that

$$l'_1 = l'_1(A_1 \mid S_0) = \int_{A_1 \times S_0} dF(a, s),$$

$$l'_2 = l'_2(A_2 \mid S_0) = \int_{A_2 \times S_0} a\, dF(a, s).$$

Note that $l'_1(\cdot \mid S_0)$ and $l'_2(\cdot \mid S_0)$ are two measures on $\gamma(A)$. Further, as F is continuous, they are non-atomic measures. It is straightforward to check that, if (l'_1, l'_2) is an element of the set of feasible tested labor inputs given S_0, then $(l'_1, l'_2) \leqslant (l'_1(\bar{A} \mid S_0), l'_2(\bar{A}^c \mid S_0))$ for some $\bar{A} \in \gamma(A)$. Hence, for fixed S_0, if for any $A_1, A_2 \in \gamma(A)$ and λ $(0 \leqslant \lambda \leqslant 1)$ there exists an $A_3 \in \gamma(A)$ such that

$$\lambda l'_1(A_1 \mid S_0) + (1-\lambda)\, l'_1(A_2 \mid S_0) = l'_1(A_3 \mid S_0) \tag{B1.a}$$

and

$$\lambda l'_1(A_1^c \mid S_0) + (1-\lambda)\, l'_2(A_2^c \mid S_0) = l'_2(A_3^c \mid S_0), \tag{B1.b}$$

then the set of feasible tested labor inputs is convex. However, Liapunov's Theorem states that if v_i $(i = 1, 2)$ are atomless measures on σ-algebra β, the set $\{(v_1(E), v_2(E)) \in R^2 \mid E \in \beta\}$ is a closed convex set (see Hildenbrand [5] for details). Hence, as $l'_1(\cdot \mid S_0)$ and $l'_2(\cdot \mid S_0)$ are atomless measures, for any given $A_1, A_2 \in \gamma(A)$ and λ $(0 \leqslant \lambda \leqslant 1)$ there exists an $A_3 \in \gamma(A)$ such that

$$\lambda l'_1(A_1 \mid S_0) + (1-\lambda)\, l'_1(A_2 \mid S_0) = l'_1(A_3 \mid S_0) \tag{B2.a}$$

and

$$\lambda l'_2(A_1 \mid S_0) + (1-\lambda)\, l'_2(A_2 \mid S_0) = l'_2(A_3 \mid S_0). \tag{B2.b}$$

Taking $l'_2(A \mid S_0)$ from both sides of (B2.b) and multiplying both sides by minus one yields

$$\lambda l'_1(A_1^c \mid S_0) + (1-\lambda)\, l'_2(A_2^c \mid S_0) = l'_2(A_3^c \mid S_0). \tag{B3}$$

Inspection of (B2.a), (B3), and the equations of (B1) yields the result that the set of feasible tested labor inputs is convex. A similar proof can be used to establish that the set of feasible untested labor inputs is convex when it is noted that, if $S_1, S_2 \in \gamma(S_0^c)$,

$$l_1^u(S_1 \mid S_0) = \int_{A \times S_1} dF(a, s)$$

and

$$l_2^u(S_2 \mid S_0) = \int_{A \times S_2} a\, dF(a, s)$$

define two atomless measures on the Borel sets of S_0^c. Hence, the convexity of the set of feasible tested labor inputs and the set of feasible untested labor inputs imply $\Theta(S_0)$ is convex. The compactness of the set $\Theta(S_0)$ follows directly from the assumptions made about the distribution function F. Thus the claims made have been established.

REFERENCES

1. G. A. AKERLOF, The market for 'lemons': qualitative uncertainty and the market mechanism, *Quart, J. Econ.* **84** (1970), 488–500.
2. K. J. ARROW, Higher education as a filter, *J. Public Econ.* **2** (1973), 193–216.
3. K. BURDETT, The testing and sorting functions of higher education, *J. Public Econ.* **10** (1978), 117–122.
4. K. BURDETT AND D. T. MORTENSEN, The existence of competitive job assignment, mimeo, Northwestern University, 1979.
5. W. HILDENBRAND, "Core and Equilibria of a Large Economy," Princeton Univ. Press, Princeton, N.J., 1974.
6. J. G. RILEY, Information, screening, and human capital, *Amer. Econ. Rev.* **66** (1976), 254–260.
7. M. ROTHSCHILD AND J. E. STIGLITZ, Increasing risk. I. A definition, *J. Econ. Theory* **2** (1970), 225–243.
8. M. ROTHSCHILD AND J. E. STIGLITZ, Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, *Quart. J. Econ.*
9. S. ROSEN, Labor quality, the demand for skills, and market selection, N.B.E.R. working paper No. 162 (1977).
10. S. ROSEN, Substitution and the division of labor, *Economica* **45** (1978), 235–250.
11. A. D. ROY. The distribution of earnings and individual output, *Econ. J.* **60** (1950), 489–505.
12. M. SATTINGER, Comparative advantage and the distribution of earnings and abilities, *Econometrica* **43** (1975), 455–468.
13. M. SPENCE, Job market signalling, *Quart. J. Econ.* **87** (1973), 355–379.
14. M. SPENCE, "Market Signalling: *Information Transfer in Hiring and Related Processess*," Harvard Univ. Press, Cambridge, Mass., 1973.
15. M. SPENCE, Competitive and optimal responses to signals, *J. Econ. Theory* **7** (1975), 296–332.

16. J. E. STIGLITZ, "Information and Economic Analysis," Technical Report No. 155, *Institute for Math. Studies, Stanford University*, 1975.
17. J. E. STIGLITZ, The theory of 'screening,' education, and the distribution of income, *Amer. Econ. Rev.* **65** (1975), 283–300.
18. C. A. WILSON, A model of insurance markets with incomplete information, *J. Econ. Theory* **16** (1977), 167–209.

JOURNAL OF ECONOMIC THEORY 25, 67-91 (1981)

Second Best Taxation as a Game

ROGER GUESNERIE

CEPREMAP and CNRS, Paris, France

AND

CLAUDE ODDOU

Centre d'Économie Quantitative, FEA, Université d'Aix-Marseille III, France

Received December 10, 1979; revised June 10, 1980

I. INTRODUCTION

Second best theory usually lies in the realm of normative analysis. It focuses attention on the "*best*" solutions which still can be obtained in problems where constraints of various natures forbid the attainment of first best Pareto optimal situations. Its central concern is the characterization of such "*second best*" solutions, and a large part of the literature concentrates on the optimal design of taxes or of policies of public firms.

Taking a second best problem, we try in this paper to shift the emphasis from "*normative*" to "*positive*" considerations. The solution we are interested in is not optimal with respect to some a priori given social welfare function embodying justice objectives; but it is supposed to reflect the power of the different agents in a negotiation process. Clearly, the conceptual tools required for such an analysis are to be found in game theory and it is actually a game theoretical approach that we follow here.

The model we are looking at is presented in Section II. It formalizes one of the simplest second best problems that can be imagined. We have a two-good economy with one private good, and one public good. Agents have different wealth in private good and the public good is financed through a wealth tax. This latter fiscal system is clearly not flexible enough to adjust contributions with marginal willingnesses to pay, so that we actually face a second best situation. There is one decision variable in the society, the tax rate which determines the public good production, about which the agents have conflicting desires. To this problem, we associate a game which provides a precise framework for the analysis of the power of the agents. The "*positive*" outcomes of the negotiation process are supposed to belong to the core of this game.

Sections III and IV concentrate on the analysis of the core. The results strongly contrast with those obtained in the classical studies of the core of

0022-0531/81/040067-25$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

economies with public good in a first best context (see, for example, Champsaur (1975)). There is no reason for the core to be large. Furthermore, in order to assure non emptiness, conditions are required reflecting for example, that the agents have similar enough opinions on the tax choice or that agents who disagree strongly with the average opinion do not own enough resources to have credible threats. Tax rates corresponding to intermediate opinions equally far from the extreme ones are shown to be more likely in the core.

We also recognize that the game under consideration has very unusual features, when compared with the classical economic games known from the literature. In particular, it is not necessarily superadditive; when people from two disjoint coalitions have too divergent feelings concerning the tax system, it may not be desirable in terms of economic efficiency that these coalitions merge putting all their resources in common. As a consequence, the core stricto sensu is empty in many economies without pathological features. For that case, we define in Section V a concept of stable structure which can be considered as an extension of the core concept to non superadditive games. This concept describes a situation where the grand coalition breaks down but where some partition, which is stable in a strong sense, emerges. This latter fact can be related with the theory of local public goods which has its origin in Tiebout (1956) and which has recently attracted a growing interest from theorists (Pestieau (1979), Stiglitz (1977), Wooders (1978), etc.). In fact, as in Westhof (1977), one can interpret the stable structure as an affectation of agents between different cities where the public good is locally produced. When the "*megalopolis*" (the grand coalition) is non viable, there still may emerge a group of different cities which define a stable arrangement in the society. Section V consists of a tentative exploration of this latter point, extending a previous attempt of Guesnerie and Oddou (1979).

II. The Public Good Game

A. *The Basic Framework*

We consider a simple economy with the following characteristics. There is one pure public good and one private good. There are n consumers indexed by $i \in N =^{\text{def}} \{1, \dots, n\}$. Consumer i has preferences on R^2_+, represented by a strictly quasi-concave utility function u_i: $u_i(x, q)$ is the utility level associated with the consumption of x units of private good and q units of public good. A constant returns to scale technology permit the transformation of one unit of private good into one unit of public good. Initial endowments are only in private goods and are privately owned. Initially, consumer i owns ω_i units of private good ($\omega_i > 0$). There are no a priori restrictions either on the distribution of endowments or on preferences (except for strict convexity).

Besides these intrinsic characteristics, we introduce in the description of our economy a basic *institutional assumption*: the public good can only be financed through a linear wealth tax. The positive tax rate t, $t \in [0, 1]$, is constrained to be the same for all individuals; it permits the withdrawal of an amount of resources $t(\sum_{i=1}^{n} \omega_i)$ to produce $t(\sum_{i=1}^{n} \omega_i)$ units of public good. The quantity of private good remaining to individual i is hence $(1-t)\,\omega_i$.

The tax system so defined is not flexible enough to adjust the contributions of individuals in proportion to their marginal willingnesses to pay, as would be desirable to meet the Samuelson rules; the Lindhal equilibrium is attainable only in exceptionnal cases. Furthermore, any direct compensating transfer between agents is ruled out. From a normative point of view, we can only expect second best Pareto optimal allocations. The derivation of tax systems maximizing some a priori given social welfare function is outside the scope of this paper but the reader is invited to check that the characterization of the Rawlsian optimum (in case of identical utility functions) is particularly straightforward. As announced in the introduction, and in constrast with the optimal taxation tradition, we are interest in deriving a positive theory which relates the decision concerning the tax system with the power of the agents in the society. For that, it is natural to associate with our problem, a game.

The formalization of the game rests on three assumptions.

First, the technology is available to any coalition, each of which can therefore transform one unit of private good into one unit of public good.

Second, the coalition alone cannot benefit from the public good produced by other coalitions. This assumption which assures that the game is orthogonal has always been made in the game theoretic study of public goods in a first best context. It is usually justified on the grounds that it reflects the maximin threat of a coalition. It will also fit well a subsequent interpretation (Section IV), according to which agents split into cities (which can be thought of as geographically distinct) so that the pure public good is locally produced in a context "*à la Tiebout*" (cf. the Introduction).

Third, institutionnal constraints on tax schemes are the same within each coalition where the public good is financed through a linear wealth tax, the tax rate being the unique decision variable.

Considering jointly these three assumptions which make our framework closely similar to that of Westhof (1977), we are in position to deduce the *characteristic form of the game*. For any $S \subset N$,

$$v(S) =$$

$$\left\{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in R^n_+,\ \exists t \in [0, 1],\ \forall i \in S,\ \alpha_i \leqslant u_i\Big[(1-t)\,\omega_i,\ t \sum_{i \in S} \omega_i\Big]\right\}.^{1}$$

---

[1] Following the notation of Shapley (1972), we associate with an agent who does not belong to the coalition any positive utility level.

Before going further, we will give a first insight into this game in a simple case.

B. *A First Insight into the Model*

Consider two agents A and B endowed respectively with a and b units of the private good. The feasible states of the coalition $\{A, B\}$ are depicted in Fig. 1. When the tax rate is t, the consumption bundles of A and B are $E = ((1-t)a, q)$ and $F = ((1-t)b, q)$, their utility levels are $u(E)$ and $u(F)$. Suppose that A is a dictator in the coalition in the sense that he can choose the tax rate without modifying the basic tax law. Figure 1 indicates that he would choose the consumption bundle G corresponding to q_A^* units of public good and then obtains an utility level u_A^*. B as a dictator would choose H associated with q_B^* units of public good and an utility level u_B^*.

Hence, Fig. 2, which depicts the set of Pareto optimal feasible utilities for coalition $\{A, B\}$, is obtained when q varies in $[q_A^*, q_B^*]$.

Let us consider now what A can do when he is alone. From Fig. 3, he will refuse any tax rate for the coalition $\{A, B\}$ which would lead to a level of

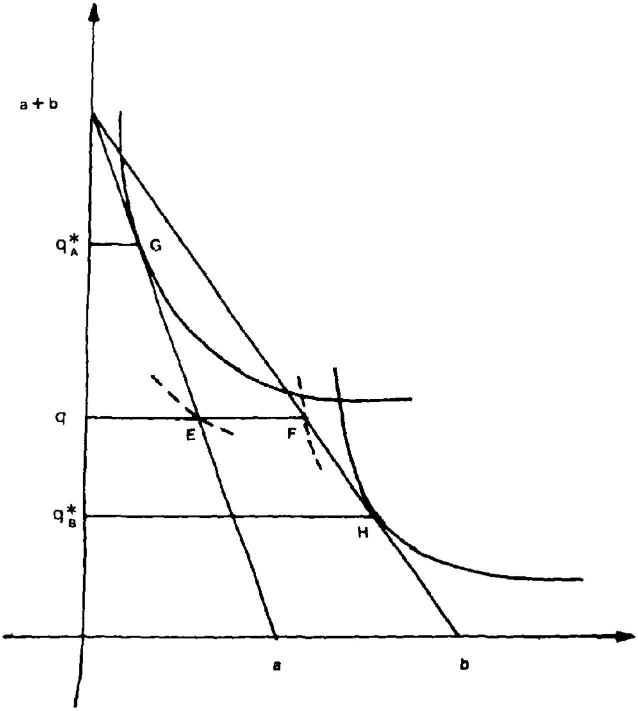


FIGURE 1

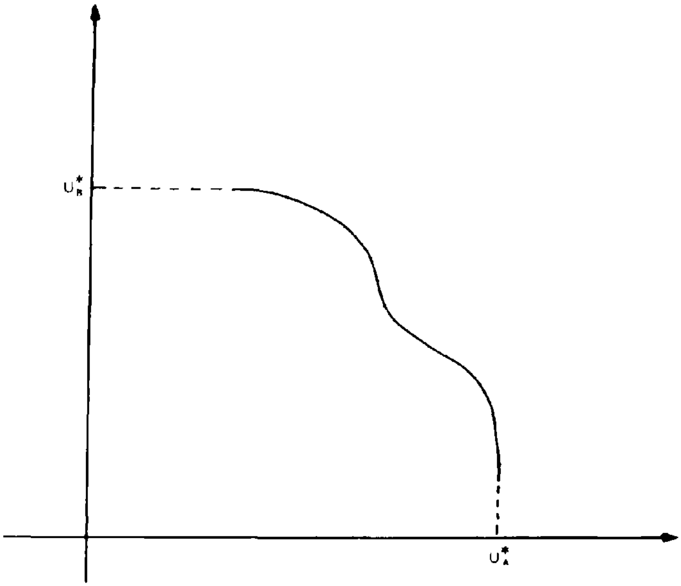


FIGURE 2

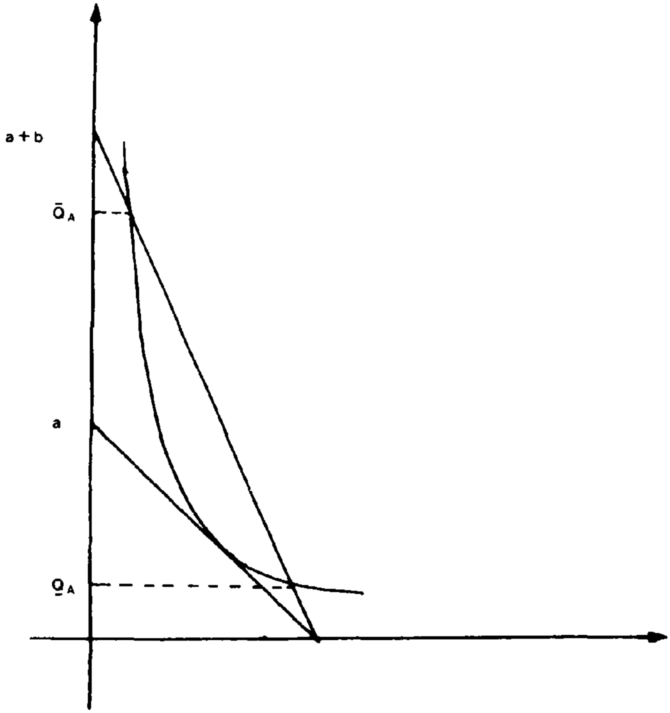


FIGURE 3

public good outside $\Delta_A = [\underline{Q}_A, \bar{Q}_A]$ since such public good levels give him an utility level smaller than $\bar{u}_A$.

Similarly B will refuse any level of public good outside $\Delta_B = [\underline{Q}_B, \bar{Q}_B]$ which would give him an utility level less than $\bar{u}_B$. If the intersection of Δ_A and Δ_B is empty (and the reader will easily convince himself from the diagram that this may actually occur) A and B will have no interest in cooperating.

Figure 4 illustrates several facts which are crucial for the analysis.

(α) The efficient points of the grand coalition do not necessarily coïncide with the efficient points of the game. In other words, from the efficiency point of view, it may not be desirable that the grand coalition form. It is equivalent here to say that the game v is not necessarily superadditive, a property which strongly contrasts this game with most games previously studied in the economic literature.

(β) The core stricto sensu, i.e., the set of allocations attainable by the grand coalition and blocked by no coalition, may be empty. In this simple two-person game, efficiency through the grand coalition, superadditivity and non emptines of the core obtain simultaneously when they are in general independent properties. However, they have connections which will be

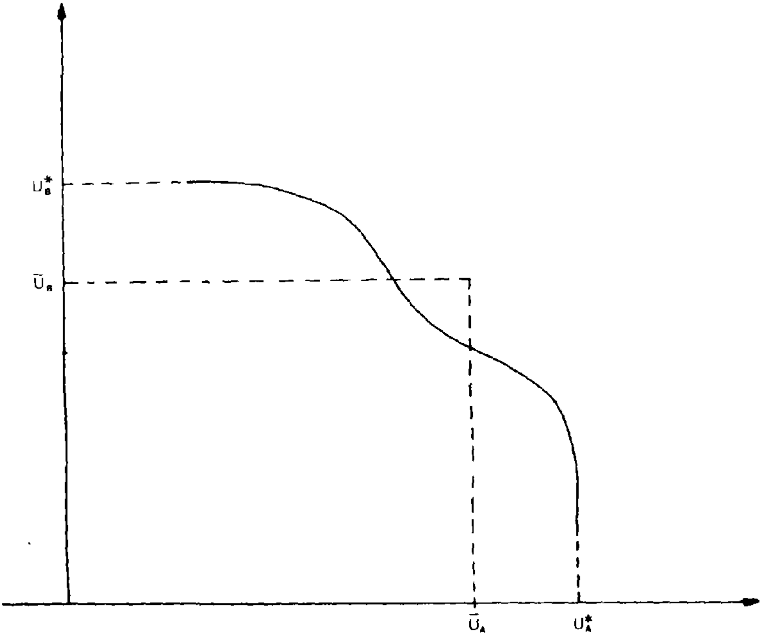


FIGURE 4

established in Section III devoted to the study of the core of superadditive games. Section IV will consider the core of non necessarily superadditive games.

(γ) Although the core is empty, a stable arrangement consisting of A and B in isolation emerges from Fig. 4. In the language of the theory of local public goods, the grand city made of A and B will split and two "*small*" cities A and B will emerge each one producing locally its own public good. The case where the core is empty and the emergence of stable structures, which can be interpreted as independent cities, is analyzed in Section V.

III. SUPERADDITIVITY, EFFICIENCY OF THE GRAND COALITION AND EXISTENCE OF THE CORE

We have just noticed that the formation of the grand coalition was not necessarily desirable from the efficiency point of view. In order to take into account this fact more precisely, we introduce the following concepts:

A structure $\mathscr{S} = (S_1, S_2, \ldots, S_p)$ is a partition of the set of agents into distinct coalitions $S_1, S_2, \ldots, S_p$.

An efficient outcome of the game is a vector $\bar{u} \in R^n_+$ such that:

(1) There is a structure $\mathscr{S} = (S_k)_{k \in K}$ for which $\bar{u} \in \bigcap_{k \in K} v(S_k)$.
(2) For all structure $\mathscr{S}' = (S_l)_{l \in L}$, $\{u \in R^n_+, u \gg \bar{u}\} \cap [\bigcap_{l \in L} v(S_l)] = \phi$.[2]

An efficient structure $\mathscr{S} = (S_k)_{k \in K}$ is a structure such that there is $\bar{u} \in \bigcap_{k \in K} v(S_k)$, which is an efficient outcome.

A structure $\mathscr{S}$ is *universally efficient* if any efficient outcome $\bar{u}$ belongs to $\bigcap_{S \in \mathscr{S}} v(S)$.

In other words, when a structure is efficient, it should be implemented when some specific arbitrage between consumers welfare prevails, i.e., for some individualistic social welfare function. A structure is universally efficient if its implementation is desirable whatever the social welfare function chosen or whatever the specific arbitrage scheme between consumers welfare. The reader is invited to illustrate these definitions with the example of the preceding section. He will also notice that in the model[3] of this paper, the structure constituted by the grand coalition alone, $\{N\}$, is always efficient, so that $\{N\}$ is the only possible candidate for universal efficiency. Let us also notice that when the game is superadditive suppose that it is superadditive. In that case, the simple example of subsection II.B suggests that, loosely speaking, the agents do not have too divergent opinions concerning the tax system. As an illustration of this fact, consider two disjoint coalitions P, T ($P \cap T = \emptyset$) and a couple of agents $i \in P, j \in T$. Suppose that i and j respectively obtained their dictatorial utility levels $u_i^*(P)$, $u_j^*(T)$. It is straightforward that superadditivity requires that $u_i^*(P)$, $u_j^*(T)$ are feasible in $S = P \cup T$. In other words, i and j being dictators in P and T would find an agreement for merging P and T favourable to each of them.

---

[2] $u \gg v \Leftrightarrow u_i > v_i, \forall i$.

[3] But not necessarily in other models. See, for example, Shaked (1978).

Formally, this can be expressed as condition (α):

Condition (α) *for coalition* S. *Existence of bilateral merging agreements.* For all P and T such that $P \cup T = S$ and $P \cap T = \emptyset$, for all $i \in P$, for all $j \in T$, there is $t \in [0, 1]$ such that

$$u_i^*(P) \leqslant u_i(t, S), \qquad u_j^*(T) \leqslant u_j(t, S). \tag{3.4}$$

Equivalently,

$$\Delta(i, u_i^*(P), S) \cap \Delta(j, u_j^*(T), S) \neq \emptyset. \tag{3.5}$$

Condition (α) can still be explained in economic terms as follows: take two agents, i and j, belonging to two distinct coalitions; suppose that these two agents have the power to decide, in their self interest, on whether the two coalitions to which they belong should merge and on which conditions; condition (α) says that such a bilateral merging agreement is always possible, whatever the couple of coalitions, whatever the couple of agents, and whatever their initial situations in the coalitions (and particularly if each of them had a dictatorial utility level).

It is remarkable that condition (α), which is necessary for superadditivity, is also sufficient.

PROPOSITION I. *The public good game is superadditive if and only if condition* (α) *holds for any coalition* S.

Proof. To prove that condition (α) is sufficient, we have to show that for any S and for any two-partition of S, $\{P, T\}$, we have

$$v(S) \supset v(P) \cap v(T).$$

This is true if and only if for any $\bar{v} \in v(P) \cap v(T)$,

$$\bigcap_{i \in S} \Delta(i, \bar{v}_i, S) \neq \emptyset. \tag{3.6}$$

But when a finite number of intervals is such that any two of them have a

As we already noticed in Section II, the property is straightforward in the case of $n=2$. For $n=3$ and $n=4$, the proof proceeds by inspection: balanced families are examined one by one and the proof that they are Scarf balanced rests on ad hoc specific arguments.

However, we were not able to extend the proof further. Moreover, it is very likely that public good games may not be Scarf balanced from $n=6$.

However, a more direct argument not relying on the Scarf condition shows that public good games which are superadditive have actually a non empty core, even if they may not be Scarf balanced. More generally, we have:

PROPOSITION IV. *If N is universally efficient, then the public good game has a non empty core.*

Proof. Let $q \in [q_1^*, q_n^*]$ be the set of second best Pareto optimal levels of public good.

Let Q^- be the set of $q \in [q_1^*, q_n^*]$ which are blocked by some coalition $P \subset I(q)$.

Let Q^+ be the set of $q \in [q_1^*, q_n^*]$ which are blocked by some coalition $T \subset J(q)$.

Our definition of blocking referring to strict inequalities, the sets Q^+ and Q^- are open in $[q_1^*, q_n^*]$. Moreover Q^+ and Q^- are intervals of the form $Q^- =]q^-, q_n^*]$, $Q^+ = [q_1^*, q^+[$.

This comes from the fact that $q \in Q^-$ (resp. $q \in Q^+$) implies $[q, q_n^*] \subset Q^-$ (resp. $[q_1^*, q] \subset Q^+$).

Let us now state:

Assertion I. For any $q \in [q_1^*, q_n^*]$, there is no blocking coalition $S = P \cup T$ such that $\varnothing \neq P \subset I(q)$, $\varnothing \neq T \subset J(q)$ and no blocking coalition contains an i such that $q_i^* = q$.

Assertion II. There is no $q \in [q_1^*, q_n^*]$ simultaneously blocked by two coalitions $P \subset I(q)$ and $T \subset J(q)$.

If both assertions were true, one would have $q^- \geqslant q^+$ (from Assertion II) and any q in $[q^-, q^+] \neq \varnothing$ would be (from Assertion I) in the core of the game.

It remains to prove Assertions I and II.

(α) If S is a blocking coalition, there exists a q^0 such that for every $i \in S$, $u_i(q^0, S) > u_i(q, N)$. But if $i \in I(q)$, necessarily $q^0 < q$ and if $i \in J(q)$, $q^0 > q$; the two are simultaneously impossible, which proves the first part of assertion I. The second part of the assertion is trivial.

(β) Suppose the contrary. Let $\bar{u}_i$ be the utility levels of the agents $i \in P \cup T$ in the corresponding blocking coalitions. From universal efficiency, there is $\bar{q}$ s.t. $u_i(\bar{q}, N) \geqslant \bar{u}_i > u_i(q, N)$, $\forall i \in P \cup T$. But this is impossible, since from q, one cannot increase in N the welfare of two agents having a q_i^* on each side of q. ∎

Proposition IV has two corollaries. The first one already has been mentioned.

COROLLARY I. *If the public good game is superadditive, then it has a non empty core.*

The second one combines Proposition II and IV.

COROLLARY II. *If condition (α) concerning the existence of bilateral merging agreement holds for the grand coalition N, then the public good game has a non empty core.*

This last statement is particularly interesting since he gives a criterion for the non emptiness of the core simple enough to be checked through a procedure based on a finite number of experiments as those defined at the end of Section III.A.

In summary, the main conclusions to be drawn from this section are the following:

When the formation of the grand coalition is certainly desirable, in terms of efficiency (certainly in the sense that it is not contingent on the specific social welfare arbitrage which is made), *it can be implemented in a stable way*; i.e., there exists a choice of tax rate which is unblocked.

The property of universal efficiency of the grand coalition, which guarantees the non emptines of the core, is itself equivalent to a simple condition which reflects that the agents do not have too divergent opinions on the tax system, and which relates to the existence of the set of potential bilateral merging agreements as expressed in condition (α).

IV. THE CORE OF THE PUBLIC GOOD GAME WHEN IT IS NOT NECESSARILY SUPERADDITIVE

We will still focus in this section on the study of the core of the public good game, but with an approach different in two respects from the approach of Section III.

The field covered is distinct: the results apply to a class of games which are not necessarily superadditive; it does not include or is included in the class of superadditive games.

The nature of the results differs: while in Section III, we only established existence theorems, the propositions given in this section either exhibit some public good level which is in the core (Proposition V) or give some interval public good levels where one can find at least one element in the core (Proposition VI).

A. *Testing whether a "Median" Level of Public Good is in the Core*

The proof of Proposition IV suggests that good candidates for the core are public good levels which are in an intermediate average position between q_1^* and q_n^*. We will consider here in particular a resource-weighted median level of public good defined as follows:

q^m *is a resource-weighted median level of public good* if an only if

$$\begin{aligned}&\sum_{i\in I(q)} \omega_i \leqslant \tfrac{1}{2} \sum_{i\in N} \omega_i \overset{\text{def}}{=} \tfrac{1}{2}\omega(N),\\ &\sum_{i\in J(q)} \omega_i \leqslant \tfrac{1}{2}\omega(N).\end{aligned} \tag{4.1}$$

Inequalities (4.1) define one q^m which is unique but for exceptional cases.

In the following we will denote $\Gamma^-(q)=\sum_{i\in I(q)}\omega_i$ (resp. $\Gamma^+(q)=\sum_{i\in J(q)}\omega_i$) the total resources of agents who desire less (resp. more) public good than q (in the sense that their dictatorial level in the grand coalition is greater or smaller than q).

In the definition of q^m, (4.1) can be written

$$\Gamma^-(q^m)\leqslant \tfrac{1}{2}\omega(N), \qquad \Gamma^+(q^m)\leqslant \tfrac{1}{2}\omega(N). \tag{4.2}$$

Loosely speaking, q^m is a R.W.M. level of public good if both people who want more than q^m and people who want less than q^m in the grand coalition own less than half of total resources.

We will now give a criterion assuring that q^m is in the core of our public good game. This criterion rests on the analysis of the answer of the agents of the economy to a simple question about which a poll could be organized in the society. The question is the following:

"Suppose that you have to choose between accepting q^m in N or being a dictator in a coalition which owns some fraction α of the total resources of the society $\omega(N)$. From which minimum level $\bar{\alpha}$ would you prefer the second solution?"

Let $\bar{\alpha}_i$ be the answer of consumer i. Supposing that individual wealths ω_i are known, for every $\gamma\in[0,1]$ one can infer from the answers $\bar{\alpha}_i$ the proportion of total wealth owned by people whose answer is smaller than γ.

Let $\beta(\gamma, q^m)$ be this number.

Formally, if we denote

$$A(\gamma, q^m)=\{i\in N, u_i^*(\gamma\omega(N))>u_i(q^m, N)\}, \tag{4.3}$$

where $u_i^*(\lambda) = \mathrm{Max}_{t \in [0,1]}\, u_i((1-t)\,\omega_i, \lambda t)$,[5]

$$\beta(\gamma, q^m)\,\omega(N) = \sum_{i \in A(\gamma, q^m)} \omega_i. \tag{4.4}$$

If, furthermore, the dictatorial level of agent i in the grand coalition were known, one could deduce from the poll

$$\begin{aligned} \bar{A}(\gamma, q^m) &= \{i \in J(q^m), u_i^*(\gamma\omega(N)) > u_i(q^m, N)\}, \\ \underline{A}(\gamma, q^m) &= \{i \in I(q^m), u_i^*(\gamma\omega(N)) > u_i(q^m, N)\}, \end{aligned} \tag{4.5}$$

and

$$\bar{\beta}(\gamma, q^m)\,\omega(N) = \sum_{i \in \bar{A}(\gamma, q^m)} \omega_i, \qquad \underline{\beta}(\gamma, q^m)\,\omega(N) = \sum_{i \in \underline{A}(\gamma, q^m)} \omega_i, \tag{4.6}$$

$\bar{\beta}$, $\underline{\beta}$ and $\beta = \bar{\beta} + \underline{\beta}$ are non decreasing step functions of γ.

In order to find some coalition blocking q^m, it is necessary to find some number γ sucht that either $\boldsymbol{\beta}(\gamma, q^m) > \gamma$ or $\bar{\beta}(\gamma, q^m) > \gamma$.

We immediately obtain:

Proposition V. *If the following conditions,*

$$\underline{\beta}(\gamma, q^m) \leqslant \gamma \qquad \textit{and} \qquad \bar{\beta}(\gamma, q^m) \leqslant \gamma \quad \textit{for all } \gamma \in \left]0, \tfrac{1}{2}\right], \tag{4.7}$$

hold, then q^m, the weighted-resource median level of public good, belongs to the core.

Particularly (4.7) *is implied by*

$$\beta(\gamma, q^m) \leqslant \gamma \qquad \textit{for all } \gamma \in \left]0, \tfrac{1}{2}\right]. \tag{4.8}$$

Clearly, the fact that we have taken q^m rather than another public good level does not play a crucial role in the analysis. For any q belonging to $[q_1^*, q_n^*]$ the criterion of Proposition V would only be slightly modified (the interval where γ varies being greater) and the method proposed to check it would remain valid.

Let us also remark that as the indirect preferences of agents are here single peaked, q^m the R.W.M. level is the *Condorcet* winner of a majority voting procedure when all agents have the same resources (or when votes are distributed in proportion of wealth). Hence, it comes out:

---

[5] Actually, we defined elsewhere $u_i^*(S)$. But there is no risk of confusion between the two notations and $u_i^*(S) = u_i^*(\omega(S))$ with $\omega(S) = \sum_{i \in S} \omega_i$.

COROLLARY I. *If all agents have the same resources, and if condition (4.7) or (4.8) of Proposition V is satisfied, the Condorcet winner public good level belongs to the core.*

B. *Conditions Assuring that "Intermediate" Public Good Levels Are in the Core*

In this section, we will consider some interval $[a, b]$ included in $]q_1^*, q_n^*[$, the set of second best Pareto optimal public good levels, and we will exhibit a condition assuring that some public good level in $[a, b]$ belongs to the core of the public good game.

Actually, reminding that $\Gamma^-(b) = \sum_{i \in I(b)} \omega_i$ and $\Gamma^+(a) = \sum_{i \in J(a)} \omega_i$, and defining $\Gamma = \max\{\Gamma^-(b), \Gamma^+(a)\}$, the condition we will consider can be introduced:

Condition (β) (with respect to the interval $[a, b]$) consists of the following two requirements:

$$\Delta(i, u_i^*(\Gamma), N) \cap [a, b] \neq \varnothing, \qquad \forall i \in N, \tag{4.9}$$

$$\Delta(j, u_j^*(\Gamma), N) \cap \Delta(i, u_i^*(\Gamma), N) \neq \varnothing, \qquad \forall j \in J(a), \qquad \forall i \in I(b). \tag{4.10}$$

Requirement (4.9) means that for every $i \in N$, there is $q^i \in [a, b]$ such that $u_i(q^i, N) \geqslant u_i^*(\Gamma)$; there is some public good level q^i in $[a, b]$ which is better for i when he is in N than his dictatorial level in a coalition which has an amount of resources Γ.

Requirement (4.10) means that if one considers one agent j who would desire in the grand coalition more public good than a, and an agent i who wants less public good than b, then there is $\tilde{q}$ such that $u_j(\tilde{q}, N) \geqslant u_j^*(\Gamma)$ and $u_i(\tilde{q}, N) \geqslant u_i^*(\Gamma)$. j and i being dictators in coalitions with an amount of resources Γ would find it profitable to join the grand coalition where some mutually agreed level of public good would prevail.[6]

The amount of resources, Γ, which is considered in condition (β) is the maximum of the amount of resources respectively owned by the agents who "*desire*" less public good than b and by the agents who "*want*" more public good than a. One sees that when $[a, b]$ is a very small interval around q^m, Γ

---

[6] Instead of condition (β) one could have stated condition (β') which, as the reader will easily check, is equivalent.

Condition (β'). $\forall j \in J(a)$, $\forall i \in I(b)$,

$$[a, b] \cap \Delta(j, u_j^*(\Gamma), N) \cap \Delta(i, u_i^*(\Gamma), N) \neq \varnothing.$$

In condition (β'), the requirement corresponding to (4.9) is suppressed; in counterpart, j and i should agree on a public good level in N which belongs to $[a, b]$.

is approximatively equal to $\omega(N)/2$; it increases when the size of the interval increases. For that reason, condition (4.10) is more likely to be satisfied when the size of the interval is small; in counterpart, condition (4.9) is more likely to be satisfied when the interval is large. It may be expected that the most interesting statement will be obtained for an intermediate size of the interval.

For example, an intermediate $[a, b]$ is obtained through the following choice: $a = q_{(1/3)}$, $b = q_{(2/3)}$, where $q_{(1/3)}$ and $q_{(2/3)}$ are defined by

$$\sum_{i \in J(q_{(2/3)})} \omega_i \leqslant \tfrac{2}{3}\omega(N), \qquad \sum_{i \in I(q_{(1/3)})} \omega_i \leqslant \tfrac{1}{3}\omega(N).$$

$$\sum_{i \in I(q_{(2/3)})} \omega_i \leqslant \tfrac{2}{3}\omega(N), \qquad \sum_{i \in J(q_{(1/3)})} \omega_i \leqslant \tfrac{1}{3}\omega(N).$$

It should be noticed that $q_{1/3}$ and $q_{2/3}$, as defined here, are in general (but not always), unique. Although the definition may look rather complicated, it is designed such that the three groups of agents—those who want less than $q_{(1/3)}$, those who want more than $q_{(2/3)}$ and those who have a dictatorial public good level between $q_{(1/3)}$ and $q_{(2/3)}$—have roughly one third of total resources; $[q_{(1/3)}, q_{(2/3)}]$ is then a kind of mediam interval in $[q_1^*, q_n^*]$.

Instead of condition (β), one introduces

$$\Delta(i, u_i^*(\tfrac{2}{3}\omega(N)), N) \cap [q_{(1/3)}, q_{(2/3)}] \neq \emptyset \cdot \forall i. \tag{4.11}$$

$$\Delta(i, u_i^*(\tfrac{2}{3}\omega(N)), N) \cap \Delta(j, u_i^*(\tfrac{2}{3}\omega(N)), N) \neq \emptyset, \forall i, \forall j. \tag{4.12}$$

We are now in position to give two additional results.

PROPOSITION VI. *If for an interval* $[a, b]$, *condition* (β) (*or condition* (β')) *is satisfied, then the public good game has a non empty core which has a non empty intersection with* $[a, b]$.

COROLLARY VI.1. *Suppose that*

(β1) *Every agent prefers some* $q \in [q_{(1/3)}, q_{(2/3)}]$ *in the grand coalition to being a dictator in a coalition owning* $\frac{2}{3}$ *of total resources* (*condition* (4.11)).

(β2) *Any couple of agents who are dictators in coalitions, each of which has* $\frac{2}{3}$ *of total resources, would agree on some common level of public good in the grand coalition* (*condition* (4.12)).

Then, the public good game has a non empty core with some element in $[q_{(1/3)}, q_{(2/3)}]$.

The proof of Proposition VI will be sketched later on.[7] Let us, however, remark now that among the criteria assuring the non emptiness of the core which have been exhibited here, Corollary VI.1 is one of the most attractive. Particularly, the number of experiments (in the sense of Section III) required for its verification is (when n is large) of the order of magnitude of n^2, which has to be compared with 2^n, the order of magnitude associated with the criterion of universal efficiency. In counterpart, the criterion is somewhat more demanding, since the two agents considered in ($\beta 2$) own together $\frac{4}{3}\omega(N)$ and not $\omega(N)$ as in condition (α) of Section III.

Still Corollary VI.1 can be considered as another illustration of the idea according to which the existence of the core relates with convergent (or not too much divergent) opinions of the agents concerning the tax system.

Proof of Proposition VI (*Sketch*). We associate with v the following game W:

$$W(S) = v(S) \qquad \text{if} \quad S \text{ is a subset of } I(b) \text{ or of } J(a),$$

$$W(N) = \{v \in R^n_+ \ \exists q \in [a, b], v_i \subseteq u_i(q, N), \forall i \in N\},$$

$$W(S) = \{v \in R^n_+, v_i = 0, \forall i \in S\} \qquad \text{for other } S.$$

We first prove that W has a non empty core: considering $\mathscr{C}$ an balanced family and $\bar{u} \in \bigcap_{S \in \mathscr{C}} W(S)$, we have to prove that $\bar{u} \in W(N)$ or equivalently $[\bigcap_{l \in N} \Delta(1, \bar{u}_1, N)] \cap [a, b] \neq \varnothing$. Or still equivalently (Helly's theorem in one dimension):

$$\Delta(1, \bar{u}_1, N) \cap \Delta(k, \bar{u}_k, N) \cap [a, b] \neq \varnothing, \quad \forall (l, k) \in N^2. \tag{4.13}$$

For $1 \in I(b)$, $k \in J(a)$ this is proved from (4.10) in condition (β) and the property of inclusion of segments Δ (cf. (3.3)). For $1 \notin I(b)$, $k \notin I(b)$ (resp. $1 \notin J(a)$, $k \notin J(a)$), a similar argument applies with (4.9).

The game W is Scarf balanced and hence has a non empty core. We show then that $\bar{q}$ (in $[a, b]$), which is in the core of W, is in the core of v (blocking coalitions should be subsets either of $I(\bar{q})$ or of $J(\bar{q})$ and hence of $I(b)$ or $I(a)$). ∎

V. The Public Good Game when the Core Is Empty

When the core is empty, two different directions of reflection are open. The first one leads to examination of restrictions on the formation of coalitions which would guarantee again the existence of the core (Section

---

[7] For Corollary VI, it is enough to note that (4.11), (4.12) do imply condition (β) for $[q_{(1/3)}, q_{(2/3)}]$.

V.A); the second one turns the attention to other possible solutions; so in Section V.B, we will discuss the possibility of emergence of *stable structures* associated with *stable solutions*, a generalization of the concept of core.

A. *Restrictions on the Formation of Coalitions*

Let us suppose that we associate with each consumer i a number $\theta_i \in [0, 1]$ with $\sum_{i=1}^{n} \theta_i = 1$. ($\theta_i$ could be interpreted as a number of votes and could, for example, be chosen such that $\theta_i = 1/n$.)

We will first restrict blocking coalitions to the family

$$\mathcal{S} = \left\{ S \subset N, \sum_{i \in S} \theta_i > \tfrac{1}{2} \right\}.$$

Then, we have:

PROPOSITION VII. *If the blocking coalitions are restricted to belong to the family* $\mathcal{S}$ *defined above then the core is non empty.*

Proof. Let us come back to the proof of Proposition III and particularly consider Assertion I.

Take a $\bar{q}$ defined by: $\sum_{i \in I(\bar{q})} \theta_i \leqslant \frac{1}{2}$ and $\sum_{i \in J(\bar{q})} \theta_i \leqslant \frac{1}{2}$ (such a $\bar{q}$ exists but is not necessarily unique).

It is left to the reader to show that such a $\bar{q}$ is in the core of the restricted game.

An obvious corollary obtains:

COROLLARY VII.1. *If blocking coalitions are restricted to coalitions with strictly*[8] *more than $n/2$ members, then the core is non empty.*

In some sense, the above analysis emphasizes that in the public good game, coalitions of small size have an important blocking power.

We should also mention here that neither in Proposition VII nor in Corollary VII.1 have we proved that the game restricted to the particular family of blocking coalitions was Scarf balanced.

In fact we were only able to prove that the public good game was Scarf balanced when restricted to blocking coalitions of $n-1$ persons.[9] As previously seen in Section III, Scarf's condition is not very useful for proving non emptiness.

---

[8] The strict inequality in the definition of S is necessary: consider the example of a two agents economy of Section III with an empty core.

[9] In this case the condition is trivially satisfied for all balanced families but $\Phi = (1, 2, \ldots, n-1\}, \{2, 3, \ldots, n\}, \{3, 4, \ldots, n, 1\}, \ldots)$. To prove Scarf balancedness we consider q^S such that $\forall i \in S$ $\bar{u}_i = u_i(q^S, S)$ [where $\bar{u}_i$ is the utility level attained in the balanced family]. Considering $\bar{\bar{q}}_i = \min q^S$, $\bar{q}_i = \max q^S$, we show by using Helly's theorem that the segments $[\bar{\bar{q}}_i, \bar{q}_i]$ have a non empty intersection and thus conclude.

We will finally mention, for the sake of completeness, another problem suggested by the classical studies on the core of economies with many agents. Consider an r-replica of our original economy; the perunit cost of the public good in the grand coalition decreases steadily. For that reason, in a first best context, the core of the public good economy becomes very large. One could have expected that in our second best context, the core would have ceased to be empty for a large enough value of r.

Actually, this is wrong:

PROPOSITION VIII. *There are economics such that whatever r, the r replica of the public good game of Section* II *has an empty core.*

We will briefly describe an example where such a phenomenon occurs.

The agents have the same initial endowments:

$$\omega_{ij} = 1, \qquad i = 1, 2; \qquad j = 1, \ldots, r.$$

Preferences are Cobbb–Douglas:

$$u_1(x, y) = x^{\alpha} y^{\beta},$$
$$u_2(x, y) = x^{\beta} y^{\alpha} \qquad (\alpha + \beta = 1).$$

For a given t, agents of type 1 and 2 respectively obtain

$$u_1(t) = r^{\beta} 2^{\beta} t^{\beta} (1 - t)^{\alpha},$$
$$u_2(t) = r^{\alpha} 2^{\alpha} t^{\alpha} (1 - t)^{\beta}.$$

The coalition consisting of all agents of type 1 (resp. 2) can guarantee to its members $\bar{u}_1 = r^{\beta} \beta^{\beta} \alpha^{\alpha}$ (resp. $\bar{u}_2 = r^{\alpha} \beta^{\beta} \alpha^{\alpha}$).

The tax rate t belongs to the core, $\forall r$, if

$$t \in [\alpha, \beta], \qquad g(t) = \left(\frac{2t}{\beta}\right)^{\beta} \left(\frac{1-t}{\alpha}\right)^{\alpha} > 1,$$

$$f(t) = \left(\frac{2t}{\alpha}\right)^{\alpha} \left(\frac{1-t}{\beta}\right)^{\beta} > 1.$$

For α small enough, these two inequalities are contradictory; the core is empty whatever r. ∎

B. *Existence of Stable Structures*

When the core is empty no agreement on the tax rate, which meets the stability requirement of being unblocked, can be expected within the grand coalition. There is a tendency for the agents to split. However, as noted in Section II, there might emerge a partition of N, with different groups of agents with different tax rates and different levels of public good, which meets stability requirements very similar to those underlying the concept of core. We will introduce here the concept of stable structure (see the concept of structure introduced at the beginning of Section III) which formalizes this idea (see also Guesnerie and Oddou (1979)).

A C-stable solution is a vector $u^0 \in R^n_+$ which satisfies the following properties:

(i) There is $\mathscr{S} = (S_k)_{k \in K}$ such that $u^0 \in \bigcap_{k \in K} v(S_k)$, where $\mathscr{S}$ is a structure;

(ii) for all $S \subseteq N$, $v(S) \cap \{u \in R^n_+, u \gg u^0\} = \varnothing$.

A stable structure is a structure $\mathscr{S}$ such that there is a stable solution u^0 belonging to $\bigcap_{S \in \mathscr{S}} v(S)$.

Several straightforward remarks are in order:

A C-stable solution is necessarily an efficient outcome.

An utility vector in the core is a C-stable solution associated with the stable structure made of $\{N\}$ alone.

If N is universally efficient, the only possible stable structure is $\{N\}$ itself.

This latter remark makes clear that the concept of stable solution is useless for superadditive games but suggests that it might reveal appropriate for our problem.

Let us briefly emphasize the conceptual connections between the notions of core and of C-stable solution. In some sense, the C-stable solution generalizes the concept of core, but it can also be viewed as merely an adaptation of the core concept.[10] To make this remark precise; let us define:

(1) $\tilde{\bar{v}}$: the smallest superadditive game associated with v

$$\tilde{\bar{v}}(S) \overset{\text{def}}{=} \bigcup_{\mathscr{S} \in \mathscr{P}(S)} \left[\bigcap_{T \in \mathscr{S}} v(T) \right],$$

where $\mathscr{P}(S)$ is the set of all partitions of S.

(2) $\tilde{v}$: an intermediate game between v and $\tilde{\bar{v}}$

$$\tilde{v}(S) = v(S) \qquad \text{if} \quad S \neq N \quad \text{and} \quad \tilde{v}(N) = \tilde{\bar{v}}(N).$$

The interest of these definitions relies in the following properties, proved in Guesnerie and Oddou (1979):

v has a strongly stable solution if and only if $\tilde{\bar{v}}$ has a non empty core; $\tilde{\bar{v}}$ has a non empty core if and only if $\tilde{v}$ has a non empty core.

---

[10] Particularly, Shaked (1978) uses a similar concept without distinguishing it from the concept of core, but this appears natural in the context of his model.

Let us now comment the meaning of the notion of C-stable solution and stable structure in our public good game.

A stable structure defines a partition of agents between different subsets. Different tax rates prevail in different groups; one group can benefit from its own public good but not of the public good produced by the other groups. The outcome cannot be blocked by the grand coalition (since it is efficient) nor by any other coalition.

Actually, what is here a structure is exactly analogous to a set of cities in Westhof (1977), and a stable structure can be interpreted as a stable structure of different cities.[11] As in Westhof, the exclusion of use of the public good between cities could be justified by the nature of the public good or by spatial considerations. It has anyway to be considered as an interesting polar case in the theory of local public goods.

What can be said about the existence of stable structure?

A first result can be given:

PROPOSITION IX. *For $m < 4$, the public good game has a C-stable solution.*

For $m = 4$, if all agents have the same resources, there is a C-stable solution.

The proof can be found in Guesnerie and Oddou (1979), where it is shown that in these cases, $\bar{\bar{v}}$ is Scarf balanced.

We will establish an extension of this last result for economies with different types of agents. For that, let us consider an economy $\mathscr{E}$ with n types of agents and r_i agents in each type. To the economy $\mathscr{E}$ let us associate the economy $\tilde{\mathscr{E}}$ with n agents: each agent i represents a type, has the initial resources $r_i \omega_i$, and hs a utility function $\tilde{u}_i(x, y) = u_i(x/r_i, y)$, where ω_i and u_i respectively are the initial resources and preferences of an agent of type i in $\mathscr{E}$.

We can prove the following lemma:

LEMMA. *Consider the public good game. If for the economy $\tilde{\mathscr{E}}$ there exists a stable structure then the same obtains for the economy $\mathscr{E}$.*

Proof. We first remark that in any C-stable solution in $\mathscr{E}$, two agents of the same type obtain the same utility level (straightforward). Then, we notice that if a coalition S gives to its members an utility vector u_i, $i \in S$, a coalition gathering agents of S and agents who are not in S but who are of a type represented in S, can give to its members of type i more than u_i. Hence, we can restrict ourselves to consider in $\mathscr{E}$ blocking coalitions which contain all the agents of the types which are represented in S. It follows that a

---

[11] Actually, our stability requirements are considerably stronger than those of Westhof.

C-stable solution in $\tilde{\mathscr{E}}$ corresponds a C-stable solution in $\mathscr{E}$ since any blocking coalition in $\mathscr{E}$ would result in a blocking coalition in $\tilde{\mathscr{E}}$.

It follows:

PROPOSITION X. *If the economy consists of three different types of agents, the public good game has a stable structure.*

If the economy consists of four different types of agents, the total resources of each type being the same, then the public good game has a C-stable structure.

We will give a final statement concerning the existence of stable structures. This statement does not rely on basically new arguments but takes advantage of the knowledge we have already acquired.

PROPOSITION XI. *Let* $\mathscr{S} = (S_k)_{k \in K}$ *be a structure which has the following properties:*

(1) *The subgame associated with the subset of agents who are in* S_k *has* S_k *as an universally efficient structure.*

(2) *Any subgame associated with a subset of agents which do not coincide with one of the* S_k *has an empty core.*

Then, φ *is a stable structure.*

Proof. From Proposition III, the subgame associated with S_k has a non empty core. Let q_k be a public good level belonging to the core of this subgame. We will prove that the sequence of $(\ldots, q_k, \ldots)$ defines a C-stable solution.

Suppose that this allocation is blocked by some coalition $S^{(0)}$, for some public good level $q^{(0)}$. Because of (1), $S^{(0)}$ cannot be a subset of some S_k. But from (2), $q^{(0)}$ in $S^{(0)}$ is blocked by some subcoalition $S^{(1)} \subset S^{(0)}$ associated with some $q^{(1)}$. But $S^{(1)}$ cannot be a subset of some $S_k \cdots$. Then, the argument can be repeated for $S^{(2)}$, $q^{2)}, \ldots, S^{(p)}, q^{(p)} \cdots$. As p cannot tend to infinity, a contradiction obtains. ▮ Q.E.D.

Proposition XI can be expressed in another way. Let us say that a coalition is *internally stable* if the associated subgame has a non empty core. *Then a structure is stable if its elements are the only internally stable groups in the society.* Such a condition is clearly rather restrictive.

VI. CONCLUSION

Let us first mention some possible extensions of the analysis presented in this paper.

As mentioned in Footnote [4], several public goods could have been incorporated in the analysis of superadditivity and universal efficiency, and actual generalizations of the corresponding results could have been obtained without major difficulties. Actually, Helly's theorem provides a guideline for extension of most of the statements where it intervenes.

The reader will also have noticed that the assumption that the wealth tax is linear only serves the purpose of assuring the connectedness of the sets $\Delta(i, q, S)$. Many results of the paper would remain true with a non linear tax schedule, under the condition that it belongs to some appropriate class of functions and that it only depends on one parameter.

On the other hand, the extension of the analysis to an economy with several private goods, where, for example, the wealth tax would be replaced by transaction taxes would raise difficult problems. In that case, the set of feasible states would have a much more complicated mathematical structure.

Finally, let us notice some open problems. Whether the existence of stable structures in the sense of Section V is "frequent" or, contrarily, unlikely is still unclear and we have not even been able to exhibit a counterexample to the existence of stable structures (a counterexample which at least requires four agents or five agents with the same initial resources).

On the other hand, the assumption that "coalitions" or "cities" do not cooperate once the public good has been produced (or do not take into account this possible cooperation before hand), can only be considered as formalizing a polar case. Its relaxation would raise the difficult conceptual problems associated with non orthogonal games.

ACKNOWLEDGMENTS

The authors are indebted to K. Binmore, and A. Shaked for stimulating conversations.

REFERENCES

BERGE, C. (1959), "Espaces topologiques. Fonctions multivoques," Dunod, Paris.

CHAMPSAUR, P. (1975), How to share the cost of a public good, *Internat. J. Game Theory* **4**, 113–129.

GUESNERIE, R. AND ODDOU, C. (1979), On economic games which are not necessarily superadditive. Solution concepts and application to a local public good problem with few agents, *Econ. Lett.* **3**, 301–306.

PESTIEAU, P. (1979), The optimality of the Tiebout model, *in* "The Political Economy of Fiscal Federalism" (Woates, Ed.), Lexington Books, Lexington, Mass.

SHAKED, A. (1978), Human Environment as a Local Public Good, London School of Economics, London.

SHAPLEY, L. S. (1972), "On Balanced Games without Side Payments," The Rand Corporation, Santa Monica, Calif. 4910.

STIGLITZ, J. E. (1977), The theory of local public goods, *in* "Economics of Public Services" (M. Feldstein and R. Inman, Eds.), pp. 214–233, Macmillan Co., New York.

TIEBOUT, M. (1956). A pure theory of local expenditures, *J. Pol. Econ.* **64**, 413–424.

WESTHOF, F. (1975), Existence of equilibria in economies with a local public good, *J. Econ. Theory* **14**, 84–112.

WOODERS, M. (1978), Equilibria, the core and juridiction structures in economics with a local public good, *J. Econ. Theory* **18**, 328–348.

JOURNAL OF ECONOMIC THEORY 25, 92–100 (1981)

Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox

ROBERT W. ROSENTHAL

Bell Telephone Laboratories, Murray Hill, New Jersey 07974

Received April 8, 1980; revised July 2, 1980

1. INTRODUCTION

The thesis of this paper is that finite, noncooperative games possessing both complete and perfect information ought to be treated like one-player decision problems. That is, players ought to assign at every move subjective probabilities to every subsequent choice in the game and ought to make decisions via backward induction. This view is in contrast with the game-theoretic approach of Nash equilibrium.

After expanding on this view for games in the abstract in Sections 2 and 3, attention is turned in Section 4 to an example due to Reinhard Selten, called the chain-store paradox, which possesses the flavor of a situation involving a predatory-pricing monopolist. It is argued that for the chain-store game the decision-analytic approach leads, under certain assumptions, to more realistic outcomes than the standard Nash-equilibrium approach.

2. AN ILLUSTRATIVE EXAMPLE

Consider the game tree in Fig. 1 to be played only once in which the outcomes are assumed to be expressed in U.S. dollars and x and y are dollar values known to both players. In words, if Player 1 chooses Left, he receives x dollars and 2 receives y dollars; if Player 1 chooses Right, then the outcome is either (0, 0) or (1 million, 1) depending on Player 2's choice. Assuming that each player's von Neumann–Morgenstern utilities for the outcomes are ordered in the same way as the dollar values and assuming that the game is played under conditions of complete information (i.e., each player knows the rules, the dollar payoffs for both, von Neumann–Morgenstern utility images of both players' payoffs, the fact that the other player knows all of this, the fact that the other knows that he knows, etc.) what should Player 1 do? More directly, for what values of x and y would you as Player 1 be indifferent between your two choices? For

0022-0531/81/040092-09$02.00/0

[illegible] Press, Inc.

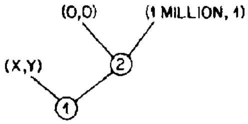


FIG. 1. Example 1.

most of those to whom I have posed this question, the answer depends on (among other things) what is known about Player 2. When I reply that Player 2 is known to be a very sensible person whose identity will always remain secret and to whom Player 1's identity will always remain secret, the answer invariably involves some x-value strictly less than 1 million.

I interpret this answer to mean that when the stakes are sufficiently high people will usually not act as though the hypothesis that other people are utility maximizers is absolutely dependable. Moves of other players simply can never be anticipated with certainty. If that is the view of real game players, what are its implications for the theory of games in which nearly all of the models and solution concepts seem to be based on the idea that all players can be assumed to be acting in their own best interests?

One might respond that the hypothesis of complete information is too strong to be ever realized in life, that utility payoffs of others (for example) can never be known with certainty and that 1's selection of Left over Right when $x < 1$ million simply indicates a natural equilibrium strategy in the imagined incomplete-information game. My feeling is that while probably no games are played under conditions of complete information in the real world, there are some situations in which the information is complete enough that the situations can reasonably be modeled as complete-information games; and that for any extensive-form model it should be possible to imagine that and reason as though the model corresponded closely enough to some hypothetical game situation. In my view, therefore, appealing to the lack of realism of the complete-information game would not be sufficient to explain preferences of Left over Right in Example 1.

It might also be pointed out that the Left strategy for Player 1 is in fact part of a Nash equilibrium whenever $x \geqslant 0$ and that game theory therefore admits considerations of the aforementioned sort. But 1's Left is only in equilibrium against 2's Left strategy; a situation I interpret to mean that 1 believes 2's (possibly nonexplicit) threat to play Left, a threat which can only be believed if 2 is possibly not a utility maximizer. Note that this strategy combination is in equilibrium whenever $x \geqslant 0$ and is otherwise insensitive to the magnitude of x. (Nash equilibria containing threats which do not get carried out are examples of what are called imperfect equilibria. In some recent game-theoretic work it has been suggested that attention should focus only on perfect equilibria or special subsets of the set of perfect equilibria. See [3, 7, 8].)

The point of view which I shall adopt is that Player 1 should consider his Right strategy to be a risky one and that the risk should be treated in the same way that it would be if Player 2 were replaced by chance in this game. That is, Player 1 should arrive at some subjective probability distribution over 2's strategy choice and should play his Right strategy if the expected utility to him from the lottery exceeds his certain utility for his Left strategy. Moreover, I will argue in Section 3 that such an approach is appropriate for the analysis of all finite games with perfect information. Note that the uncertainty should concern Player 2's choice, not his utilities for the various outcomes. Although this distinction is not significant in Example 1, where 1's positing appropriate uncertainty over Player 2's payoffs and assuming he does not carry out threats can be seen to be equivalent to 1's positing uncertainty over 2's choice, the distinction does become significant in multi-move games.

3. Perfect-Information Games

A finite n-person extensive-form game with perfect information consists of:

1. A tree with a distinguished node (called the origin). Node b follows a if there is a path from the origin to b through a. A terminal node is a node with no followers.
2. A partition of the nonterminal nodes into sets labeled 0, 1, 2,..., n. A node in set $i > 0$ corresponds to a move by Player i. A node in set 0 is a move by chance.
3. For every node in set 0, a probability distribution over the immediate followers of that node. The interpretation should be clear.
4. For every terminal node, a vector in R^n. The ith component of the vector is interpreted as Player i's von Neumann–Morgenstern utility for the outcome represented by the terminal node.

For games with perfect information, no extra informational complications arise. When a player is to choose an immediate follower from any of his nodes, he knows for which node he is making the choice. Thus, he knows perfectly all the choices which have been made at preceding nodes. There are no possibilities of secret or simultaneous moves. The assumption of complete information is that everything about the extensive form is known to all the players of the game.

A pure strategy for Player i (>0) is a function which assigns to each of Player i's nodes one of its immediate followers. A Nash equilibrium in pure strategies is a pure strategy combination (one strategy for each player) with the property that no player can improve his expected utility payoff by switching to some other pure strategy if the strategies of all other players are held fixed. It is well known (see, e.g., [1]) that every finite perfect-information game possesses at least one Nash equilibrium in pure strategies and that such an equilibrium can always be constructed by working backward through the tree, selecting a best (defined inductively) immediate follower at every personal node and taking expected utilities at every chance node. Nash equilibria obtained in this way will be termed principal equilibria. Nonprincipal pure-strategy equilibria may exist, as has already been observed in Example 1. Nash equilibria in randomized strategies (i.e., probability distributions over pure strategies), which have been widely studied, are of lesser interest for games with perfect information; since there seem to be no intuitive reasons to randomize consciously when the outcome of the randomization will be evident to all players who have subsequent moves.

Although a player may not consciously intend to randomize and may in fact intend to reason inductively in order to select his strategy in a game with perfect information, it may be that for some reason his intention is not realized. If players with preceding moves realize that this may be the case, they would be wise to consider it when making their choices.

Why might players believe that players moving subsequently might deviate from their principal-equilibrium strategies? (Incidentally, I do not rule out the possibility that a player may view his own subsequent choices as random.) One possibility is a mistake. Another is that subsequent players may decide not to analyze the tree completely because the time and effort required are not justified by the potential gains. (In this case it could be argued that the game in question is only part of the more complicated problem in which decisions about time and effort are also made. Such considerations quickly lead to unmanageably large models and even questions about infinite regressions. It seems far simpler to study the game of perfect information with appropriate allowance made for moves which are not part of principal equilibria.) A third possibility (though related to the previous one) is that the player in question recognizes that the game is actually played under conditions of incomplete information, but that a model incorporating the uncertainties would be unmanageable for him. To keep matters simple, he assumes a game with complete information but modifies his principal-equilibrium calculations in a manner he feels to be intuitively appropriate. Still a fourth possibility is the belief that other players may be acting according to some other view of what is rational. This view may not known with certainty.

Consider Example 2 (see Fig. 2), where the origin is at the left. If Player 2 is actually called upon to move in this game, he has clear evidence that Player 1 is not playing according to any Nash-equilibrium strategy. Should he ignore this evidence, or should he consider the possibility that Player 1

4. Predatory Pricing and the Chain-Store Paradox

Consider a single-product industry containing a single firm setting a monopoly price and making corresponding profits. If another firm enters the industry, there is enough business for both to survive, but profits for the monopolist will be severely cut. A possible strategy for the monopolist is to cut price whenever a new firm enters to a level low enough that the entrant will sustain losses in every period. This strategy presumably reduces the monopolist's profits even further for those periods when the low price is in effect (perhaps even causing losses for the monopolist). If the monopolist drives the entrant out, and if no new firm enters, however, the monopolist can resume his high-price, high-profit position. Price cutting to drive other firms out of an industry or to deter entry is called predatory pricing and is a subject of considerable interest to economists. The proper role of anti-trust procedures in controlling predatory pricing is a matter of some debate among both economists and law makers (see, for example, [4, 6]).

The question of whether or not and to what extent the threat of predatory pricing is likely to succeed in deterring entry under various conditions is a little-understood but important aspect of the debate. It is also a question with a strong game-theoretic flavor.

In [7], Reinhard Selten describes an example of a game, called the chain-store paradox (version one), which seems to capture the essential elements of the deterrence question. In Selten's game there is one chain store, Player A, with branches in 20 towns. In each town there is one potential entrant i ($i = 1, \ldots, 20$). Furthermore, for each i there is only one date at which he can enter. These dates are sequential with Player 1's date first, then 2's, etc. If Player i enters on his date, A with full knowledge has the choice of whether to respond aggressively or cooperatively. The payoffs to Players i and A (although it is only a partial payoff to A who is concerned with the sum of his 20 payoffs) are given below:

i's decision	A's decision in period i	i's payoff	A's payoff
In	Cooperate	2	2
In	Don't cooperate	0	0
Out	—	1	5

At Player i's entry date, he is fully aware of all previous moves.

There are two features of this game which deserve comment with respect to the predatory-pricing interpretation. One is the assignment of unique entry dates to each firm. This assumption seems to have been made to keep the specification of the game as simple as possible and does not appear to distort seriously the deterrence question. The second feature is the finite-time horizon assumption. Firms are often assumed to plan for the infinite horizon (with discounting where appropriate). Furthermore, it is well known that in many repeated games there is a qualitative difference between the structure of the set of Nash equilibria in the finitely-repeated case (even with a sufficiently large number of repetitions) and that in the infinitely-repeated case. Since all of the discussion thus far has been directed at finite-move games, it would be improper to argue that any insights gained by reasoning inductively on the finite, chain-store game are valid for a situation in which players perceive an infinite horizon. But it would also seem that there should be very little difference between what a firm would actually do early in a large-but-finite horizon predatory-pricing situation and what that same firm would do in a similar infinite-horizon situation.

The thrust of Selten's analysis of the chain-store game is that while the principal Nash equilibrium of the game requires all Players *i* to enter and A to cooperate, one's intuition suggests perhaps that the threat of the aggressive response might suffice to deter entry, at least until late in the game. Selten then proceeds to an interesting discussion of the game from many different points of view.

The point which I wish to make is that there is a considerable similarity between the chain-store game and Example 3. If the discussion of the previous sections has been convincing, then the approach advocated here for games with perfect information may be of use in resolving the "paradox" and in analyzing predatory pricing and other deterrence situations.

Let us look at the chain-store game under the assumption that the size of utility differences affects subjective probabilities in the same direction as in Example 3. At each of Player A's last nodes he gains two units of utility by picking the cooperative response. Suppose, however, that Player 20 views that choice as less than certain. At 20's move, therefore, his expected utility gain is less than one from choosing In. Player A therefore has an expected gain of less than 2 if he responds aggressively to 19's choice of In. If 19 considers the aggressive response to him to have higher probability than 20 considered the aggressive response, then he has even less to gain than 20 by choosing In. As in Example 3, the effects may quickly compound if the probabilities are appropriate monotone functions of the expected utility differences. The reader will be spared any further details. It should be obvious that for a wide class of methods for forming subjective probabilities, the outcome of the game will be that entry is deterred for a large portion of the potential entrants.

REFERENCES

1. H. W. KUHN, Extensive games and the problem of information, *in* "Contributions to the Theory of Games, Vol. II" (Kuhn and Tucker, Eds.), Annals of Mathematics Studies, No. 28, pp. 193–216, Princeton Univ. Press, Princeton, N. J.
2. R. D. LUCE AND H. RAIFFA, "Games and Decisions," Wiley, New York, 1957.
3. R. B. MYERSON, Refinements of the Nash equilibrium concept, *Internat. J. Game Theory* 7 (1978), 73–80.
4. R. A. POSNER, "Anti-Trust Law: An Economic Perspective," Chap. 8, Univ. of Chicago Press, Chicago, 1976.
5. H. RAIFFA, "Decision Analysis," Addison–Wesley, Reading, Mass., 1968.
6. F. M. SCHERER, "Industrial Market Structure and Economic Performance," Chap. 8, Rand McNally, Chicago, 1970.
7. R. SELTEN, The chain-store paradox, *Theory and Decision* **9** (1978), 127–159.
8. R. SELTEN, Reexamination of the perfectness concept for equilibrium points in extensive games, *Intern. J. Game Theory* **4** (1975), 25–55.

JOURNAL OF ECONOMIC THEORY 25, 101–130 (1981)

Length and Cycle Equalization

RONALD A. HEINER

Department of Economics, Brigham Young University,
Provo, Utah 84602

Received April 11, 1978; revised December 6, 1979

1. INTRODUCTION

One of the ways of attempting to solve the paradoxes of collective choice theory has been to equalize together options involved in cyclical intransitivities. The best known approach was axiomatized by Schwartz [19] (see also [4, 13]). These methods have certain limitations, an example being the violation of the Pareto principle. This paper investigates the general area of cycle equalization via the *length* of cycles equalized. A structure is developed which ranges from completely unrestricted length methods at one extreme (the Schwartz choice function being a particular example) to the opposite extreme of maximally restricted length just short of no cycle equalization at all.

A basic reason for considering length restriction is that unrestricted cycle equalization can be very weak in ability to distinguish between options. McKelvey [16], for example, shows in the case of a multi-dimensional option space and euclidian preferences, that majority vote will either be completely transitive with no cycles or else will engulf the entire option space in one massive cycle (thus, the Schwartz method would select everything). Length restriction is a direct way of increasing distinguishability, expecially if constrained to a maximum length which is independent of the number of options being ordered or chosen from. This is developed in Section 2.2, the major obstacle being that restriction to less than the number of options considered may eliminate the transitivity properties of the equalization process. It is thus also developed which type of cycle equalization permits length restriction while still preserving transitivity properties.

The Pareto principle is also analyzed. This has been a principal objection to unconstrained length methods (see, for example, [3; 2, p. 238]). It is seen in Section 5.3 that over the dual spectrum of type and length of cycle-equalization only one method will permit the guarantee of both the Pareto principle and acyclicity. A basic corollary is that maximal length restriction just short of no equalization is required to achieve the Pareto principle.

0022-0531/81/040101–30$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

A general notion of independence from irrelevant changes in the set of individual preferences is introduced in Section 4 and analyzed relative to length restriction in Section 5. The condition is called "Independence from Irrelevant Information" (denoted III). It is a dual condition requiring independence from certain types of changes in individual preferences for a given set of options; and from changes in the set of options for given individual preferences. The basic idea is to not allow any change in the social choice from any factor which does not involve options (by deletion or change in preferences) that would themselves be chosen or socially tied with any chosen option. This independence specification is related to Arrow's independence condition and to the weak axiom of revealed preference extended to collective choice behavior (see Arrow [2], Sen [21]).

Finally it is shown that the joint requirement of III, distinguishability, transitivity, and the Pareto principle singles out a special class of maximally length-restricted cycle equalization (denoted $\dot{R}'_2$) as the only method to simultaneously achieve all of these attributes. This typifies the difference in attainable properties between the two extremes of the length spectrum; $\dot{R}'_2$ at one end, and, for example, the Schwartz method at the other end.

2. MOTIVATION AND DEFINITIONS

2.1. *Introductory Concepts*

Let $K = \{1,..., n\}$ be the set of individuals and X be the universal set of options, with β representing the set of all finite nonempty subsets of X. It is usually assumed that individual preferences are defined over all of X. However, real world voting systems never actually consider preferences over all potential options when choosing from a particular feasible set. Indeed, preference information is typically submitted only over the set of options considered in a given situation. Thus, the domain of individual preferences is specified to permit preferences over subsets of X to be submitted. (S, r_i) is a weak ordering for $i \in K$ of options S, a subset of X. $(S, r) = (S, r_1),..., (S, r_n)$. Strict individual preferences are denoted by p_i, and defined by $xp_i y$ if and only if $xr_i y$ and not $yr_i x$. Then D is defined as the set of all (S, r) for all $S \in \beta$, and represents all potential sets of *preference data* over subsets of X. In addition, the explicit representation of both S and r enables a clear distinction between two types of changes in the set of preference information considered:

(1) The change from (S, r) to (S', r) for $S' \subset S$ means the *content*, r, of individual rankings is unchanged but now restricted to a subset S' contained in S.

(2) The change from (S, r) to (S, r') means the *scope*, S, of individual

preference data evaluated is unchanged, but now the content over a fixed S has changed from r to r'.

Distinguising between (1) and (2) is crucial to understanding the condition, Independence from Irrelevant Information, to be presented below.

The collective choice is divided into two steps of first determining a *social preference relation*, $\mathscr{P}$, and then choosing options using the maximizing function, $C(A, \mathscr{P}) = \{x \in A \mid x\mathscr{P}y \text{ for all } y \in A\}$. The collective ranking $\mathscr{P}$ is a function of $(S, r) \in D$ into a connected, binary relation $\mathscr{P}$ over S, and this functional dependence is denoted by $\mathscr{P}(S, r)$. The *collective choice function*, $C(A, \mathscr{P}(S, r))$, is defined for all $A \in \beta$ and $(S, r) \in D$ such that $A \subset S$, so that preference information is evaluated at least over the set of options chosen from.

Now consider informally the notion of cycle equalization. Suppose the collective ranking $\mathscr{P}$ implies, xPy, yPz, zPx. Such a sequence of preferences from x to x is called a cycle; in this case a 3-cycle because three options or equivalently three preferences are involved. In general, the *length* of a cycle refers to the number of distinct options or preferences involved in the cycle. The longest cycle possible for a given set of options S is simply the total number of options in S, here denoted $|S|$. For example, for $S = \{x, y, z, w\}$ with four options the longest cycle is length 4, and is one of six possible 4-cycles involving $\{x, y, z, w\}$.[1]

A convenient way of representing a cycle is by specifying a sequence of preferences which begins with some x and ends with some y. More precisely $x\hat{y}_m$ represents such a sequence of m *strict*-preferences; that is, $x\hat{y}_m = \{a_0 P a_1, \ldots, a_{m-1} P a_m\}$, $a_0 = x$, $a_m = y$, $m \geqslant 1$. For example, $x\hat{y}_1 = xPy$ and $x\hat{y}_2 = \{xPz, zPy\}$. An m-cycle is thus denoted $x\hat{x}_m$, which occurs whenever $x = y$ for some $x\hat{y}_m$. When $n \geqslant 2$, $x\hat{y}_m$ is called an *indirect comparison*, and $x\hat{y}_1 = xPy$ is called a *direct comparison*. It is immediate that the union of an indirect comparison $x\hat{y}_m$ with an "opposite" direct comparison $y\hat{x}_1$ forms a cycle of length $m + 1$. For example, $x\hat{y}_2 = \{xPz, zPy\}$ and $y\hat{x}_1 = yPx$ form the 3-cycle $\{xPz, zPy, yPx\}$.

Thus, one way a pair options x, y can be involved in a cycle is if both a direct and indirect comparison connect them in "opposite directions"; that is, both $x\hat{y}_1$ and $y\hat{x}_m$ for $m \geqslant 2$. More generally, cycles are implied whenever two indirect comparisons $x\hat{y}_m$, $y\hat{x}_n$ connect x and y in opposite directions to form the $(m + n)$-cycle, $x\hat{x}_{m+n}$.

Now, *the basic idea of cycle equalization is to rank indifferent* (*or*

---

[1] The six possible 4-cycles are: $\{xPy, yPz, zPw, wPx\}$, $\{xPz, zPy, yPw, wPx\}$, $\{xPz, zPw, wPy, yPx\}$, $\{xPy, yPw, wPz, zPx\}$, $\{xPw, wPz, zPy, yPx\}$, $\{xPw, wPy, yPz, zPx\}$. The general formula for the number of distinct cycles (of varying length) between n options, where either xPy or yPx can be used between each pair of distinct options, is $\sum_2^n n!/(n-i)!\, i$. For example, with ten options, there are 1,112,073 possible distinct cycles from length 2 to length 10.

"equal") to each other those options involved in a cycle. Given the above relationships between direct and indirect comparisons forming cycles, this can be accomplished by defining a new social ordering $\mathscr{R}'$ which implies indifference $xI'y$ whenever $x\hat{y}_n$ and $y\hat{x}_m$ both occur for some initially given social ordering $\mathscr{R}$. This is the basic structure in the cycle equalization methods to be defined. However, before continuing the notion of an indirect comparison is generalized so that indifferences in the direct comparisons which form them are permitted.

To do so $x\hat{y}_1 = xPy$ is called a *strict* preference and $x\mathscr{R}y$ is called a *weak* preference. A sequence of weak preferences $\overline{xy}_m = \{x_0\mathscr{R}a_1,\ldots, a_{m-1}\mathscr{R}a_m\}$, $a_0 = x$, $a_m = y$, $m \geqslant 1$ is then called a *weak indirect comparison*, while $x\hat{y}_m$ is a *strict indirect comparison*. In addition, $x\tilde{y}_m$ is a *semi-strict indirect comparison* and represents an intermediate version between $\overline{xy}_m$ and $x\hat{y}_m$ in which at least one of the preferences which connect x and y must be a strict preference. For $m = 2$, examples of these are, $\overline{xy}_2 = \{xIz, zIy\}$, $x\tilde{y}_2 = \{xIz, zPy\}$, and $x\hat{y}_2 = \{xPz, zPy\}$. Now using strict, semi-strict, and weak indirect comparisons, various cyclical patterns are formed which may contain indifferences. Thus, in general a cycle will refer to a sequence of weak preferences or indifferences, which starts and ends with the same option. For example, the 3-cycle $\overline{xx}_3 = \{x\mathscr{R}z,\ z\mathscr{R}w,\ w\mathscr{R}x\}$ could be $\{xPz,$ $zPw,$ $wPx\}$, $\{xIz, zIw, wPx\}$ or $\{xIz, zIw, wIx\}$. These are three of a total of eight possible cycles which might connect x to x via some pattern of strict preferences and indifferences.

2.2. *Cycle Equalization and Ability to Discriminate*

As already stated, the basic idea of cycle equalization is to judge indifferent or equivalent options involved in cyclical patterns. The basic way this is done is to check if a pair of options $\{x, y\}$ are involved in two indirect comparisons which connect them in opposite sequences from x to y and from y to x. Different methods are possible depending on what pattern of strict preferences and indifferences are permitted to form indirect comparisons between pairs of options.

Besides *type* of indirect comparison permitted, different cycle equalization methods are possible by restricting the *length* of cycles permitted to be equalized together; that is, *by setting an upper limit on the number of preferences a cycle can contain and still imply the options involved be equalized together.* A key motivation for restricting the size of cycles before equalization is permitted is to minimize the degree to which such equalization is used; because when increasingly applied more and more options are judged indifferent. The most extreme form of this would be to automatically tie together all options potentially considered.

This is a basic problem with any cycle equalization method, because the very way it resolves cyclical intransitivity is by replacing such intransitivity with indifference. Yet, this resolution ceases to be attractive if too many options are tied together in order to eliminate cyclical patterns.

Thus, the objective is to investigate how to permit cycle equalization, yet minimize its use so as to preserve as many of the strict preferences, xPy, initially present in the original ordering $\mathscr{R}$. Both *type* and *length* of cycle equalization relate to this objective.

To see what is potentially at stake, consider briefly McKelvey's [16] result on cyclical majorities. He assumes the set of options is represented by all the points in a finite dimensional euclidian space, with each person having a most preferred point (called his ideal point, X_i^*). Individual preferences are "euclidian" in that all points are ranked in order of their euclidian distance from a person's ideal point ($xr_i y$ if and only if $|x - X_i^*| \leqslant |y - X_i^*|$). McKelvey shows that unless citizens' preferences satisfy a very restrictive condition (existence of a strong total median), the entire option space is connected by one massive cycle of strict majority preferences. More precisely, for any pair of options x, y in the space there is a sequence of strict majority comparisons $x\hat{y}_m$, of some length m, which connects them. Thus, for any pair of options x, y there are opposite indirect majority comparisons $x\hat{y}_m$, $y\hat{x}_n$ in the space. This implies any type of cycle equalization will tie together all options in the space unless the length of the indirect comparison used to specify the cycle equalization process is restricted.

Therefore, the basic method used in subsequent definitions is to *restrict the maximum length of indirect comparison* (and thus the maximum length of cycle) permitted in the specification of the cycle equalization process. Now, the number of options in a set S (denoted $|S|$) itself places a bound on the maximum length of cycle or indirect comparison possible between distinct options in S. The maximum lengthed cycle is $|S|$, and the maximum lengthed indirect comparison between distinct options is $|S| - 1$. For example, if $S = \{x, y, z\}$ then the longest cycle is a 3-cycle such as $x\hat{x}_3 = \{xPy, yPz, zPx\}$, and the longest indirect comparison involves at most two preferences such as $x\hat{z}_2 = \{xPy, yPz\}$.

Thus, unless the maximum length of indirect comparison, denoted by T, is bounded to *less* than $|S| - 1$, no restriction of cycle equalization is obtained beyond that automatically implied by the size of S itself. As will be seen, the bound, $T < |S| - 1$, plays an important role in the behavior of a cycle equalization process.

3. The Cycle Equalization Methods

We are now ready to define the cycle equalization methods to be analyzed in this paper. The basic structure is to derive from a given ordering function

$\mathscr{R}(S, r)$ a new ordering function, say $\mathscr{R}^*$, which will judge x at least as preferred as $y(x\mathscr{R}^*y)$ whenever x is connected to y by a certain type of indirect comparison (either $\overline{xy}_m$, $x\tilde{y}_m$, or $x\hat{y}_m$). Thus, when two options are connected in opposite directions by two indirect comparisons, both $x\mathscr{R}^*y$ and $y\mathscr{R}^*x$ are implied, which together imply indifference xI^*y.

3.1. *Formal Definition*

To facilitate formal specification the following notation is introduced. For a given collective ordering function, $\mathscr{R}(S, r)$, $\overline{xy}_m \subset S$ means a sequence of distinct options $\{a_0, a_1, \ldots, a_m\}$ associated with some $\overline{xy}_m = \{a_0 \mathscr{R}(S, r)\, a_1, \ldots, a_{m-1} \mathscr{R}(S, r)\, a_m\}$, $a_0 = x$, $a_m = y$, $m \geqslant 1$ is contained S; that is, $\{a_0, \ldots, a_m\} \subset S$. Analogous definitions apply to $x\tilde{y}_m \subset S$ and $x\hat{y}_m \subset S$ except that now at least one of the direct preferences, or all of the direct preferences, respectively, which connect x to y must be strict preferences, $a_i P(S, r)\, a_{i+1}$.

It is again noted that the length restriction parameter T denotes the *maximum* length of indirect comparions permitted in the cycle equalization process.

We then have

DEFINITION 3.1. For a given ordering function, $\mathscr{R}(S, r)$, define $\mathscr{R}_T$, $\mathscr{R}'_T$, $\mathscr{R}''_T$, $T \geqslant 1$ by

(a)
$$\left\{\begin{array}{l} x\mathscr{R}_T(S, r)\, y \\ x\mathscr{R}'_T(S, r)\, y \\ x\mathscr{R}''_T(S, r)\, y \end{array}\right\} \quad \text{if and only } \overline{xy}_1 \subset S, \qquad \text{or}$$

$$\left\{\begin{array}{l} \overline{xy}_m \subset S \\ x\tilde{y}_m \subset S \\ x\hat{y}_m \subset S \end{array}\right\} \quad \text{for} \quad 1 \leqslant m \leqslant T.$$

(b) $x\mathscr{R}'''_T(S, r)\, y$ if and only if either $x\hat{y}_m \subset S$ for $1 \leqslant m \leqslant T$; or for no $1 \leqslant m \leqslant T$, $(x\hat{y}_m \subset S$ or $y\hat{x}_m \subset S)$.

A few remarks concerning these methods are made:

(1) The main difference between the four methods is that more stringent conditions are required for an indirect comparison to imply an option is at least as preferred as another. For $\mathscr{R}_T$, an indirect comparison composed entirely of indifferences is all that is required. For $\mathscr{R}'_T$, $\mathscr{R}''_T$ the indirect comparison must contain at least one, or contain all strict direct preferences to imply $x\mathscr{R}'_T\, y$ or $x\mathscr{R}''_T\, y$, respectively. In addition, a weak preference,

$x\mathscr{R}y=\overline{xy}_1$, is sufficient also to imply x at least as preferred as y for the first three methods. Thus, a direct indifference xIy implies indifference for $\mathscr{R}_T$, $\mathscr{R}'_T$, $\mathscr{R}''_T$, for all $T \geqslant 1$. However, a direct indifference xIy does not imply indifference for $\mathscr{R}'''_T$ unless there is no indirect sequence (of all strict preferences) connecting x to y or y to x.

The effect of these different specifications is to restrict the ability to equalize cycles, by requiring indirect comparisons contain more strict preferences. Thus, for any given T, the ability to distinguish increases over the four types of equalization from $\mathscr{R}_T$ the least discriminating to $\mathscr{R}'''_T$ the most discriminating (see Appendix A for a precise characterization).

(2) The length restriction parameter T imposes, for each given method, an additional requirement for indirect comparisons to imply an option be ranked at least as preferred as another option. Thus, as discussed above, reducing T will further increase the ability of a method *not* to generate indifferences. (See Appendix A.)

(3) Setting $T=1$ corresponds to not allowing any cycle equalization, and as intuitively should be the case, all four methods reduce down to equivalence with the ordering function from which they are derived; that is $\mathscr{R}=\mathscr{R}_1=\mathscr{R}'_1=\mathscr{R}''_1=\mathscr{R}'''_1$.

(4) If no indifferences in the direct preferences occur, the four methods are identical to each other. Thus, the difference between the methods relates essentially to how they equalize cycles not composed only of strict preferences.

(5) Finally, the four methods (or any cycle equalization process) will leave an ordering function unchanged whenever it is transitive. That is, if $\mathscr{R}$ is transitive, then $\mathscr{R}=\mathscr{R}_T=\mathscr{R}'_T=\mathscr{R}''_T=\mathscr{R}'''_T$ for all $T \geqslant 1$.

4. INDEPENDENCE FROM IRRELEVANT INFORMATION

The condition to be formally specified, Independence From Irrelevant Information (III) requires the social choice be invariant to certain types of changes in preference information evaluated. This invariance applies to two types of changes in the set of preference information evaluated, $(S, r) \in D$: changes in S for given r; and changes in r for given S.

4.1. *Motivation*

We wish to motivate III in relation to the *Weak Axiom of Revealed Preference* (WARP) and in relation to Arrow's original presentation of *Independence of Irrelevant Alternatives* (IIA). First consider WARP. The particular notational version is taken from Schwartz [20] and is initially written in the standard revealed choice literature notation, which explicitly represents only the feasible or eligible set chosen from, $A \in \beta$ (β was defined above as the set of all non-empty subsets of X).

WARP If $C(A) \not\subset B$, then $C(A - B) = C(A) - B$.

Thus, WARP is a rationality condition which requires a certain pattern of choice consistency with respect to changes in the feasible set, $A \in \beta$. One meaning of WARP applied to collective choice functions, which explicitly use preference information of individuals, is to require WARP consistency for any given set of preference data $(S, r) \in D$. Formally this is WARP′: If $C(A, \mathscr{R}(S, r)) \not\subset B$, then $C(A - B, \mathscr{R}(S, r)) = C(A, \mathscr{R}(S, r)) - B$. WARP′ is a very easy condition to satisfy and in fact is guaranteed by any common rank voting or point system such as the Borda or Plurality methods.

However, for real world voting systems, the set of preference information evaluated is rarely held fixed (which he hypothesis of WARP′ requires) as the feasible set varies. The typical case is to evaluate or submit preference data only over those options considered eligible in a given situation. In this case, both A and S vary together, rather than A changing for a fixed scope of information submitted. Therefore, we consider the achievement of WARP consistency when both the feasible set A and scope of information evaluated S vary in the same way; that is, if B is eliminated from A then B is also eliminated from S.

WARP* If $C(A, \mathscr{R}(S, r)) \not\subset B$, then $C(A - B, \mathscr{R}(S - B, r)) = C(A, \mathscr{R}(S, r)) - B$.

It is important to realize that the choice consistency pattern with respect to changes in the feasible set is identical for both WARP′ and WARP*, except that now WARP* requires this consistency pattern be maintained in the empirically likely case in which the scope of preference data submitted or evaluated varies in the same way the feasible set varies.

The next step is to formulate two weaker rationality conditions immediately implied by WARP*,

WARP1* If $C(A, \mathscr{R}(S, r)) \not\subset B$, then $C(A - B, \mathscr{R}(S - B, r)) \subset C(A, \mathscr{R}(S, r))$.

WARP2* If $C(A, \mathscr{R}(S, r)) \cap B = \varnothing$, then $C(A - B, \mathscr{R}(S - B, r)) = C(A, \mathscr{R}(S, r))$.

Two of the three sub-conditions of III relate directly to these weakened versions of WARP*, and are specified precisely in the next section. Before proceding, however, these conditions are also related to Arrow's original discussion of IIA [1, pp. 26–28].

First, it is noted that WARP2* relates directly to the Borda example on

p. 27 of [1]. This illustration chooses from $A = \{w, x, y, z\}$ by computing Borda scores from preference information submitted over $S = \{w, x, y, z\}$. Then, a non-chosen option y is deleted from A and Borda scores are recomputed after preference information concerning y is also deleted (i.e., y is deleted also from S). Thus, the two choices compared are $C(A, \mathscr{R}(S, r))$ and $C(A - \{y\}, \mathscr{R}(S - \{y\}, r))$, and since $y \notin C(A, \mathscr{R}(S, r))$, then WARP2* requires, $C(A, \mathscr{R}(S, r)) = C(A - \{y\}, \mathscr{R}(S - \{y\}, r))$. Yet the respective choices are in fact, $\{x\}$ and $\{x, z\}$.

At first glance, this type of condition might not appear related to IIA, which is formally defined for a fixed feasible set A and changing preferences; whereas the example involves a choice consistency between two different feasible sets A and $A - \{y\}$. However, it is easily proven that the consistency property WARP2* implies IIA, and more generally that IIA is equivalent to the stronger condition WARP* if $\mathscr{R}(S, r)$ is transitive (see Heiner [10] for a proof of this). Thus, there is a close connection between WARP* type consistency (in which the scope of preference information submitted varies in the same way the feasible set varies) and the formal specification of IIA, which is explicitly represented in the statement of III below. Arrow's condition IIA is defined within our choice system, $C(A, \mathscr{R}(S, r))$, by[2]:

IIA If (S, r), (S', r') for $A \subset S$, $A \subset S'$ satisfy $(A, r) = (A, r')$, then $C(A, \mathscr{R}(S, r)) = C(A, \mathscr{R}(S', r'))$.

A special case of this, which is the one typically assumed in the literature, is to allow only r to change so that we have IIA′: If (S, r) (S, r') for $A \subset S$ satisfy $(A, r) = (A, r')$ then $C(A, \mathscr{R}(S, r)) = C(A, \mathscr{R}(S, r'))$. (Note that Arrow does not explicitly define the underlying domain of preferences, which leaves unclear whether the scope of information submitted S is permitted to vary in the definition of IIA).

Now the third condition of III defined below concerns a stronger version of IIA′ which is not implied by WARP2* or WARP*. The reason for this is that IIA′ requires preferences over the whole feasible set A must be unchanged in order to imply the choice from A is also unaffected. Yet, there are many cases in which a stronger condition would seem desirable. For example, let $S = A$ have five options with the social ordering implied by $\mathcal{R}(S, r)$ being x, y, z, w, u, respectively, from first to last (so that x is chosen). Now suppose some individuals revise their judgements of w and u, but all other comparisons remain the same. It seems reasonable that the social ordering between w and u could change but much less so that the ordering between x and y could reverse so that now y is chosen. We thus, formulate a stronger version of IIA′ which requires that unless preference changes involve at least some comparison with a chosen option, the choice set must be the same. Formally this is done by defining the notion of *externally similar* preferences for a subset $B \subset S$.

---

[2] The correspondence to Arrow's condition can be seen more schematically if the variable π is used to denote an element $(S, r) \in D$, and defining π_A to mean the restriction of preference information to $A[\pi_A = (A, r)]$. Then IIA is: if π, $\pi' \in D$ satisfy $\pi_A = \pi'_A$ then $C[A, \mathscr{R}(\pi)] = C[A, \mathscr{R}(\pi')]$. However, because π is determined by both the scope of information S and the content of information r over S; then $\pi' = (S', r')$ corresponds potentially to variation in both S and r, not just S alone or r alone. Most analysts have interpreted IIA to mean variation in r alone (here called *ranking independence*, RI) and have ignored the possibility of variation in S for given r (called *scope independence*, SI) or the possibility of both S and r varying simultaneously (here called IIA). It can be shown that SI and IIA are equivalent and that SI implies RI; but the converse from RI to SI is false. These issues are discussed in Heiner [10].

DEFINITION 4.1 (External Similarity). For any $(S, r)(S, r') \in D$, if $\text{ext}[B, r, r']$ then for all $\{x, y\} \subset S$, and $i \in K$, $(\{x, y\}, r_i) \neq (\{x, y\}, r'_i)$ implies $\{x, y\} \subset S - B$.

The name external similarity is used because the only preference changes permitted must involve pairs of options *neither of which are in B*. Thus, no comparison with an option in B has been altered. Then, the stronger version of IIA′ is defined,

IIA* If $\text{ext}[C(A, \mathcal{R}(S, r)), r, r']$ for (S, r), $(S, r') \in D$, then $C(A, \mathcal{R}(S, r)) = C(A, \mathcal{R}(S, r'))$.

4.2. *Formal Specification of III*

The condition *Independence From Irrelevant Information* is a package of three properties which are direct analogs of WARP1*, WARP2*, and IIA*. The key difference is that changes in preference information evaluated must be isolated not only from the options initially chosen, *but also from options tied or indifferent to any chosen option*. To capture this additional factor, $C^*(A, \mathcal{R}(S, r))$ is defined as the set of all $x \in S$ such that $xI(S, r)\, y$ for some $y \in C(A, \mathcal{R}(S, r)$. Thus, $C^*(A, \mathcal{R}(S, r))$ includes all chosen options $C(A, \mathcal{R}(S, r))$, plus any other option in S which is socially indifferent to some chosen option. Then III is defined,

DEFINITION 4.2 (Independence from Irrelevant Information). 1. For all $(S, r) \in D, A, B \subset S$

(a) If $C(A, \mathcal{R}(S, r)) \not\subset B$, then $C(A - B, \mathcal{R}(S - B, r)) \subset C^*(A, \mathcal{R}(S, r))$.

(b) If $C^*(A, \mathcal{R}(S, r)) \cap B = \varnothing$, then $C(A - B, \mathcal{R}(S - B, r)) = C(A, \mathcal{R}(S, r))$.

2. If $\mathrm{ext}[C^*(A, \mathscr{R}(S, r)),\ r,\ r']$ for (S, r), $(S, r') \in D$, then $C(A, \mathscr{R}(S, R)) = C(A, \mathscr{R}(S, r'))$.

Conditions III1(a,b) relate directly to the choice consistency patterns of WARP1* and WARP2*. The key to remember is hat these rationality properties are maintained when the scope of information S varies in the same way as the set of eligible or choosable options, A, varies. III(a) concerns what would happen to the choice set if some (but not all) options which would have been selected were deleted from consideration, *also assuming preference information relative to them would not be considered either*. The answer is that the new choice set can contain only those remaining options which would have been chosen or which were indifferent to at least some of those options which would have been chosen. III1(b) extends this by requiring what would happen if the deleted options *and associated preference information* did not involve options which would have been chosen, or indifferent to some option which would have been chosen. The answer is that the choice set must be unchanged.

Condition III2 is the same as IIA* above except that again options tied with some chosen option must be considered. It says that if the *content* of individual preferences change, (for a given scope of preference information evaluated) but no preference change involves any chosen option or any option which is indifferent to some chosen option; then the choice set is unaffected. Or stated in the converse, the only way preference changes (S, r) to (S, r') can effect the choice set is that citizens must change their minds concerning some options in $C^*(A, \mathscr{R}(S, r))$.

Conditions III1 and III2 are actually different aspects of a *single* condition which requires the choice set be invarient to changes in the set of preference information evaluated (either by deletion or addition to the scope of information, or by changes in the content of a given scope of information) unless preference data concerning chosen options or options indifferent to chosen options is modified.

4.3. III, IIA, *and Purely Binary Preferences*

Having formally defined III, we wish to show a further link with Arrow's IIA via the notion of purely binary preferences, denoted $\dot{\mathscr{R}}$. Given any ordering function $\mathscr{R}(S, r)$, its *purely binary preferences* $\dot{\mathscr{R}}$ are defined by, $x \dot{\mathscr{R}}(S, r)\, y$ if and only if $x \mathscr{R}(\{x, y\}, r)\, y$; with $\dot{P}$ and $\dot{I}$ correspondingly defined $x\dot{P}(\{x, y\}, r)\, y$ if and only if $xP(\{x, y\}, r)\, y$, and $x\dot{I}(S, r)\, y$ if and only if $xI(\{x, y\}, r)\, y$. Thus, an ordering function's purely binary preferences limit the scope of preference information used *to only the pair of options directly involved* in a direct social comparison $x \mathscr{R} y$. That is, the scope of preference information evaluated is limited to only binary subsets $\{x, y\} \subset S$ when determining the corresponding social judgements $x \mathscr{R} y$, between pairs x, $y \in S$.

In addition, two ordering functions $\mathscr{R}(S, r)$ and $\mathscr{R}'(S, r)$ are said to be *companions* if and only if for all $(S, r) \in D$ and $x, y \in S$, $xP(S, r)\,y$ implies not $yP'(S, r)\,x$. Thus, companion ordering functions never have opposite preferences. We then have

THEOREM 4.1. (a) *$C(A, \mathscr{R}(S, r))$ satisfies IIA if and only if $\mathscr{R} = \dot{\mathscr{R}}$.*

(b) *If $C(A, \mathscr{R}(S, r))$ satisfies III(1a), then $\mathscr{R}$ and $\dot{\mathscr{R}}$ are companions.*

Thus, IIA is equivalent to the ordering function $\mathscr{R}$ being identical to its purely binary preferences, $\dot{\mathscr{R}}$; while III1(a) implies $\mathscr{R}$ and $\dot{\mathscr{R}}$ are companions. Besides showing another link between IIA and III, part(b) indicates that the achievement of III1(a,b) is closely related to an ordering function's purely binary preferences. To see this further, consider any ordering function $\mathscr{R}$ which is transitive, but *not* necessarily a companion to its purely binary preferences; that is, for some (S, r) and $x, y \in S$, xPy and $y\dot{P}x$. For example, the plurality and Borda methods always imply transitive orderings. However, these transitive orderings may not be companions to their purely binary preferences, which are equivalent to pairwise majorities. Also remember if $\mathscr{R}$ is transitive then a cycle equalization method derived from $\mathscr{R}$ (denoted $\mathscr{R}^*$) will leave $\mathscr{R}$ unchanged; that is, $\mathscr{R}^* = \mathscr{R}$. Now, in general, if $\mathscr{R}$ and $\dot{\mathscr{R}}$ are not companions, and $\mathscr{R}^* = \mathscr{R}$ because $\mathscr{R}$ is transitive, then $\mathscr{R}^*$ cannot be a companion to $\dot{\mathscr{R}}$. Hence, Theorem 4(b) implies no cycle equalization method $\mathscr{R}^*$ derived from $\mathscr{R}$ can satisfy III1(a).

However, a general result is still implied concerning cycle equalization and III1(a), which applies to all possible ordering functions, even to transitive ones such as Borda and plurality voting. It is obtained by applying a cycle equalizing process to the purely binary preferences of $\mathscr{R}$, rather than to $\mathscr{R}$ directly. One clue to this alternate procedure is suggested by the fact that a cycle equalization process is always a companion to the ordering function from which it is derived (because strict references xPy are never reversed, but only transformed into indifference). Hence, any cycle equalization method applied to $\dot{\mathscr{R}}$ (denoted $\dot{\mathscr{R}}^*$), is a companion to $\dot{\mathscr{R}}$, which is consistent with the implication of Theorem 4(b) that $\dot{\mathscr{R}}^*$ and $\dot{\mathscr{R}}$ must be companions for $\dot{\mathscr{R}}^*$ to satisfy III1(a). In general, it will be shown below that certain cycle equalization methods guarantee III properties when applied to the purely binary preferences, $\dot{\mathscr{R}}$, of *any* ordering function $\mathscr{R}$.

Thus, the cycle equalization methods defined above *are applied to an ordering function's purely binary preferences $\dot{\mathscr{R}}$, rather than $\mathscr{R}$ itself.* This is done by defining $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$, $\dot{\mathscr{R}}''_T$, $\dot{\mathscr{R}}'''_T$, respectively, exactly as above in Definition 3.1, except that now $\dot{\mathscr{R}}(S, r)$ is used in the definition rather than $\mathscr{R}(S, r)$; so that all of the indirect comparisons (now denoted $\dot{\overrightarrow{xy}}_m$, $x\dot{\check{y}}_m$, $x\dot{\hat{y}}_m$ are composed of only purely binary strict preferences or indifferences.

For example, $\dot{\mathscr{R}}''_T$ is defined: $x\dot{\mathscr{R}}''_T(S, r)\,y$ if and only if $\dot{\overrightarrow{xy}}_1 \subset S$ or $x\dot{\hat{y}}_m \subset S$

for some $1 \leqslant m \leqslant T$, where $x\dot{\hat{y}}_m$ contains only purely binary strict preferences.

It should also be noted that the purely binary preferences $\dot{\mathscr{R}}$ of $\mathscr{R}$ are more likely to be cyclical than $\mathscr{R}$ itself. For example, if $\mathscr{R}$ is computed from Borda scores, it is automatically transitive; yet its purely binary preferences (which are equivalent to pairwise majority votes) can be cyclical. Thus, even if an ordering function $\mathscr{R}$ is transitive, cycles may still need to be dealt with if III1(a) is achieved by applying a cycle equalization process to $\dot{\mathscr{R}}$, which can still be cyclical.

5. Main Results across the Length Spectrum

We are now prepared to consider the structure of interrelations implied across the four methods $\mathscr{R}_T$, $\mathscr{R}'_T$, $\mathscr{R}''_T$, $\mathscr{R}'''_T$, and variation in the length restriction parameter T. In the process, a special class of cycle equalization is identified, which achieves the most attractive simultaneous combination of the properties discussed. In particular, it is the only one to permit the Pareto principle and not allow strict cycles, $x\hat{x}_m$.

5.1. *The Length Spectrum*

The structure of properties associated with the cycle equalization methods can be conveniently represented by partitioning the range of T into four zones which define a spectrum from no cycle equalization at one extreme to unrestricted cycle equalization at the other. The four zones defined are:

(1) $T = 1$	This corresponds to not allowing any cycle equalization at all, because the cycle equalization methods are all identical to the ordering function from which they are derived.
(2) $T = 2$	This represents the maximum degree of length restriction permitted without eliminating the possibility cycle equalization altogether.
(3) $3 \leqslant T < \lvert S\rvert - 1$	This represents the intermediate range of cycle equalization between maximum length restriction at $T = 2$ and no restriction at $T = \lvert S\rvert - 1$. Note that $\lvert S\rvert - 1$ is the largest size indirect comparison possible between a pair of *distinct* options $x, y \in S$.
(4) $T \geqslant \lvert S\rvert - 1$	This represents unrestricted cycle equalization, because cycles involving all $\lvert S\rvert$ options in S can be equalized together. Thus, no restriction beyond that automatically implied by the size of S is required.

5.2. *The Structure of Properties across the Length Spectrum*

We are now ready to present the major structure of properties achievable across the length spectrum and across the different cycle equalization methods. Two additional properties beyond the distinguishability and independence conditions discussed above are also considered. These concern transitivity properties, and the Pareto principle.

The transitivity properties refer to full transitivity, quasi-transitivity, and acyclicity. They are equivalent to the standard definitions in the literature, except are here specified using the notation of indirect comparisons. Formally they are defined: an ordering function $\mathscr{R}(S, r)$ is *transitive*, *quasi-transitive*, and *acyclical* if and only if, respectively, for all $(S, r) \in D$, $x, y \in S$ and $m \geqslant 2$: $\overline{xy}_m$ implies $\overline{xy}_1$; $x\hat{y}_m$ implies $x\hat{y}_1$; and not $x\hat{x}_m$. Thus, an ordering function is transitive if its weak direct preferences are transitive; quasi-transitive if its strict direct preferences are transitive; and acyclical if no cycles composed of all strict direct preferences can occur.

The Pareto principle is the strong version in which no one worse off implies an option is at least as preferred socially as another, and some individuals better off implies a strict social preference. That is,

Definition 5.1 (Strong Pareto Principle). An ordering function satisfies the *strong Pareto principle* if and only if for all $(S, r) \in D$, $x, y \in S$:

(1) If $xr_i y$ for all i; then $x\mathscr{R}(S, r)\, y$.

(2) If $xr_i y$ for all i and $xp_j y$ for some j; then $xP(S, r)\, y$.

In order to present compactly and as schematically as possible, the set of results relative to the independence, transitivity, and Pareto properties are shown below in a diagram which divides the length spectrum into four zones as discussed above in Section 5.1.

An "X" in a particular box indicates a particular property is satisfied by the associated cycle equalization method, for T bounded within one of the four zones; ranging from no cycle equalization at $T = 1$, to unrestricted cycle equalization for $T \geqslant |S| - 1$. Otherwise, that property is not implied if a box is left blank.

As discussed above concerning III properties, conditions III1(a, b) cannot be guaranteed unless the cycle equalization methods are applied to an ordering function's purely binary preferences $\dot{\mathscr{R}}$. Thus, the results given pertain to cycle equalization of $\dot{\mathscr{R}}$. Certain of the results apply more generally and are specified precisely in Appendix A.

The following remarks discuss the structure of results shown in Fig. 1.

(a) *Relation to the Schwartz Method.* As shown on the diagram, $\dot{\mathscr{R}}'''_T$, is equivalent to the Schwartz choice function [denoted $C_{Sch}(A, \mathscr{R})$] for

LENGTH SPECTRUM

Properties	T = 1				T = 2				3 ≤ T < |S| - 1				T ≥ |S| - 1			
	$\dot{R}_T$	$\dot{R}'_T$	$\dot{R}''_T$	$\dot{R}'''_T$	$\dot{R}_T$	$\dot{R}'_T$	$\dot{R}''_T$	$\dot{R}'''_T$	$\dot{R}_T$	$\dot{R}'_T$	$\dot{R}''_T$	$\dot{R}'''_T$	$\dot{R}_T$	$\dot{R}'_T$	$\dot{R}''_T$	$\dot{R}'''_T$*
Transitive													X			
Quasi-transitive					X	X			X	X			X	X		X
Acyclical					X	X			X	X			X	X	X	X
IIA	X	X	X	X												
III1(a,b)	X	X	X	X	X	X			X	X			X	X		
III2	X	X	X	X	X	X			X				X			
Strong Pareto Principle**	X	X	X	X		X	X	X								

*Equivalent to the Schwartz method.
**Guaranteed if R also satisfies neutrality and positive association.

FIGURE 1

Let $A = S = \{x, y, z\}$ and for $(S, r) = (\{x, y, z\}, r_1, r_2, r_3, r_4)$:

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ x & x & y & y \\ z & z & x & x \\ y & y & z & z \end{array}$$

Majority vote implies $x\dot{I}y$, $y\dot{I}z$, $x\dot{P}z$. Thus, $\dot{\mathscr{R}}_2$, $\dot{\mathscr{R}}'_2$, $\dot{\mathscr{R}}''_2$, $\dot{\mathscr{R}}'''_2$ imply, respectively:

(1) $x\dot{I}_2 y, y\dot{I}_2 z, x\dot{I}_2 z,$

(2) $x\dot{I}'_2 y, y\dot{I}'_2 z, x\dot{P}'_2 z,$

(3) $x\dot{I}''_2 y, y\dot{I}''_2 z, x\dot{P}''_2 z,$

(4) $x\dot{I}'''_2 y, y\dot{I}'''_2 z, x\dot{P}'''_2 z.$

Thus, $\dot{\mathscr{R}}_2$ equalizes x and z even though x is unanimously preferred to z; but $\dot{\mathscr{R}}'_2, \dot{\mathscr{R}}''_2, \dot{\mathscr{R}}'''_2$ still order x above z. In addition, for $T \geqslant 3$ none of the cycle equalization methods, $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$, $\dot{\mathscr{R}}''_T, \dot{\mathscr{R}}'''_T$ can guarantee the Pareto principle, as seen in the next example.

Let $A = S = \{w, x, y, z\}$ and for $(S, r) = (\{w, x, y, z\}, r_1, r_2, r_3)$:

$$\begin{array}{ccc} 1 & 2 & 3 \\ w & y & z \\ x & z & w \\ y & w & x \\ z & x & y \end{array}$$

Majority vote implies, for $T \geqslant 3$, the cycle, wPx, xPy, yPz, zPw will be equalized by them all, even though w is unanimously preferred to x. Remembering that $\dot{\mathscr{R}}''_T, \dot{\mathscr{R}}'''_T$ can be cyclical for $T < |S| - 1$, implies that $\dot{\mathscr{R}}'_2$ is the *only possibility to guarantee both the Pareto principle and not be cyclical.*

5.3. *A Special Class Cycle Equalization*

The structure of results discussed above has identified $\dot{\mathscr{R}}'_2$ as the only method to satisfy transitivity, independence, and Pareto properties. The full set of results are here precisely formulated and stated.

The guarantee of the Pareto principle is implied via two additional properties on $\dot{\mathscr{R}}$. Both are standard concepts in the literature: neutrality and positive association. The neutrality property is taken directly from Sen [21, p. 72] and is defined: $\dot{\mathscr{R}}$ satisfies *neutrality* if and only if for all (S, r), $(S, r') \in D$ and $x, y, z, w, \in S$ if xr_iy if and only if zr'_iw, and yr_ix if and only if wr'_iz then $x\dot{\mathscr{R}}(S, r)y$ if and only if $z\dot{\mathscr{R}}(S, r')w$ and $y\dot{\mathscr{R}}(S, r)x$ if and only if $w\dot{\mathscr{R}}(S, r')z$. Thus, neutrality requires that if two options x and y respectively, have exactly the same relation to each other in each individual's preference for (S, r) as z and w have for (S, r'), then $\dot{\mathscr{R}}$ must also have the same relation between x and y for (S, r), as it has between z and w for (S, r'). Positive association is taken from Arrow's original formulation, and is defined: $\dot{\mathscr{R}}$ satisfies *positive association* if and only if for all (S, r), $(S, r') \in D$, x, $y \in S$; if $xp_i y$ implies $xp'_i y$ and $xr_i y$ implies $xr'_i y$ for all i, then $x\dot{P}(S, r)\, y$ implies $x\dot{P}(S, r')\, y$. Thus, to paraphrase Arrow [1, p. 25], if no individual lowers an option in relation to another option then the purely binary preference (which uses preference information only over those two options) between them must also not lower that option in relation to the other option.

The general result concerning the Pareto principle and cycle equalization can now be stated.

THEOREM 5.1. *Let $\mathscr{R}$ be any ordering function whose purely binary preferences satisfy neutrality, positive association, and the strong Pareto principle. Then, $\dot{\mathscr{R}}'_2$, $\dot{\mathscr{R}}''_2$, $\dot{\mathscr{R}}'''_2$ will also satisfy the strong Pareto principle.*

Now, from Theorem 5.1 and the above results we have the following properties attainable via the use of $\dot{\mathscr{R}}'_2$.

THEOREM 5.2. *Let $\mathscr{R}$ be any ordering function, then:*

(a) *$\dot{\mathscr{R}}'_2$ is quasi-transitive.*

(b) *$C(A, \dot{\mathscr{R}}'_2)$ satisfies* 1(a,b) *and* 2 *of III.*

(c) *$\dot{\mathscr{R}}'_2$ is more discriminating than $\dot{\mathscr{R}}_T$ and $\dot{\mathscr{R}}'_T$ for all T, S such that $T \geqslant 2$, $|S| - 1 \geqslant 2$ or $T > 2$, $|S| - 1 > 2$, respectively.*

(d) *If in addition $\dot{\mathscr{R}}$ satisfies neutrality, positive association, and the strong Pareto principle, then $\dot{\mathscr{R}}'_2$ also satisfies the strong Pareto principle.*

(e) *If in addition $\dot{\mathscr{R}}$ satisfies non-imposition and non-dictatorship (conditions* 4,5 *of Arrow* [1]), *then these properties are also implied for $\dot{\mathscr{R}}'_2$.*

Two comments are made:

1. The recursive definition of $\dot{\mathscr{R}}'_2$ within the overall length spectrum may obscure the intuitive meaning of $C(A, \dot{\mathscr{R}}'_2)$. If $\dot{\mathscr{R}}$ does not contain indifferences, this is particularly simple. $C(A, \dot{\mathscr{R}}'_2)$ is then equivalent to $\{x \in A \mid$ for all $y \in A$, either $x\dot{P}y$, or for some $z \in A$, $x\dot{P}z$, $z\dot{P}y\}$. Thus, chosen options are those which compared to any other option either beat that option directly, or beat some other option which beats that option. Longer chains of dominance then directly beating another option which directly beats a given option are not permitted in determining the final selection.

2. Theorems 5.1, 5.2 and the above results specify an interesting structure of properties over the spectrum of length restricted cycle-equalization. This

can be summarzied in the following. $\dot{\mathscr{R}} = \dot{\mathscr{R}}_1$ has the maximal ability to distinguish among all of $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$, $\dot{\mathscr{R}}''_T$, $\dot{\mathscr{R}}'''_T$ and its pairwise preferences have complete independence from any individual preference data not directly over the pair (i.e., the collective preference over any $\{x, y\}$ is totally independent of $(S - \{x, y\}, r)$). However, there is no guarantee against very complex cyclical patterns in $\dot{\mathscr{R}}$.

This can be overcome, regardless of the relative size of T and S, by $\dot{\mathscr{R}}_2$, $\dot{\mathscr{R}}'_2$, which are both quasi-transitive; retain a relatively strong ability to distinguish; and still enable the weakened independence properties of III 1(a,b), 2. Of these two, only $\dot{\mathscr{R}}'_2$ will permit the Pareto principle, and for all $T \geqslant 3$ none of $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$, $\dot{\mathscr{R}}''_T$, $\dot{\mathscr{R}}'''_T$, can preserve this guarantee. As T increases, the ability to distinguish for each type of equalization method diminishes, but both quasi-transitivity and III 1(a,b) can still be retained for $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$.

It is suggested, therefore, that given knowledge of these interrelationships; $\dot{\mathscr{R}}'_2$ enables the most effective simultaneous combination of transitivity, independence, distinguishability, and Pareto properties achievable across the spectrum of length restricted cycle equalization.

6. SUMMARY

We have developed a structure of length restricted cycle equalization which ranges from no equalization permitted at one extreme to unrestricted methods at the other. The Schwartz choice function is equivalent to maximization of one of the unrestricted length methods, $\dot{\mathscr{R}}'''_T$. Such methods have a very weak ability to discriminate as the number of options considered increases, and cannot guarantee the Pareto principle. Only two of the methods $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$ are not themselves cyclical if length restricted to less than the number of options considered.

Arrow's independence condition IIA is equivalent to the ordering function being purely binary and requiring that no cycle equalization be permitted. However, the independence conditions of III can be achieved by $\dot{\mathscr{R}}_T$, $\dot{\mathscr{R}}'_T$ so long as any cycle equalization is possible. Furthermore, the Pareto principle can be achieved only at the polar extreme of the length spectrum just short of no cycle equalization. If acylicity is also desired, only $\dot{\mathscr{R}}'_2$ can achieve both.

The general conclusion is that judged according to distinguishability, transitivity, independence, and Pareto properties, the maximally restricted end of the length spectrum has superior properties to the unrestricted end. In particular, $\dot{\mathscr{R}}'_2$ is singled out as implying the most attractive simultaneous combination of these attributes across all the cycle equalization methods and degree of length restriction investigated.

APPENDIX A

Here the notion of ability to distinguish is formalized and certain results depicted in the length spectrum diagram of Section 5.2 are precisely stated. In addition, the non-manipulability properties of $\dot{\mathscr{R}}'_2$ are characterized.

1. *Distinguishability*

The relation, $\geqslant$, between ordering functions is read "ordering function $\mathscr{R}$ is *at least as discriminating* as $\mathscr{R}'$;" and is defined $\mathscr{R} \geqslant \mathscr{R}'$ if and only if for all $(S, r) \in D$ and $x, y \in S$, $x\mathscr{R}(S, r)y$ implies $x\mathscr{R}'(S, r)\, y$. Thus, if $\mathscr{R} \geqslant \mathscr{R}'$ then xIy implies $xI'y$ but the converse may not hold so that $xI'y$ and xPy is possible. From $\geqslant$, $\mathscr{R} > \mathscr{R}'$ if and only if $\mathscr{R} \geqslant \mathscr{R}'$ and not $\mathscr{R}' \geqslant \mathscr{R}$. Also, a *fully cyclical* ordering function is one in which cycles composed of any pattern of direct strict preferences or indifferences is possible. It is noted that the purely binary preferences $\dot{\mathscr{R}}$ of an ordering function are more likely to be fully cyclical than $\mathscr{R}$ itself. The reason for introducing full cyclicality is that differences in ability to distinguish are not exhibited for all cyclical patterns in $\mathscr{R}$, and in the limit disappear if $\mathscr{R}$ is fully transitive.

THEOREM A1. *For any fully cyclical ordering function* $\mathscr{R}(S, r)$:

(a) $\mathscr{R}'''_T > \mathscr{R}''_T > \mathscr{R}'_T > \mathscr{R}_T$ *for all* $T \geqslant 2$;

(b) $\mathscr{R}'''_u > \mathscr{R}'''_T$, $\mathscr{R}''_u > \mathscr{R}''_T$, $\mathscr{R}'_u > \mathscr{R}'_T$, $\mathscr{R}_u > \mathscr{R}_T$,

for all $1 \leqslant u < |S| - 1$ *and* $u < T$.

Note, these results do not require purely binary preferences.

2. *The Schwartz Method*

The following relates $\mathscr{R}'''_T$ to the Schwartz choice function [18], which implicitly equalizes cycles of arbitrary length.

DEFINITION A1 (Schwartz Choice Function). For any ordering function $\mathscr{R}$, the *Schwartz Choice Function*, denoted $C_{\text{Sch}}(A, \mathscr{R})$, is defined by the following three axioms:

(1) For all $x \in C_{\text{Sch}}(A, \mathscr{R})$ and $y \in A - C_{\text{Sch}}(A, \mathscr{R})$; $x\mathscr{R}y$.

(2) If $B \subset C_{\text{Sch}}(A, \mathscr{R})$ and for all $x \in B$ and $y \in C_{\text{Sch}}(A, \mathscr{R}) - B$, $x\mathscr{R}y$; then for all such x, y, $y\mathscr{R}x$.

(3) *For* $B \subset A$, $B \neq \varnothing$. If for all $x \in A - B$ and $y \in B$, $y\mathscr{R}x$; then for some $y' \in B$, $y' \in C_{\text{Sch}}(A, \mathscr{R})$.

THEOREM A2. *Given any ordering function* $\mathscr{R}$: *For all* $(A, r) \in D$ *and* T *such that* $T \geqslant |A| - 1$; $C_{\text{Sch}}(A, \mathscr{R}(A, r)) = C(A, \mathscr{R}_T'''(A, r))$

This theorem is basically the result shown by Deb (see [5]) on the Schwartz choice function being equivalent to the "transitive closure" of the strict preference relation, P. However, Deb does not discuss the relative context of Schwartz's rule in the length spectrum.

It can also be shown that $\mathscr{R}_T'''$ is also equivalent in the sense of Theorem A2 to the choice function developed by Kalai *et al.* in [13]. In addition, the "majority algorithm" process and "optimal set" presented in Cambell [4] can be directly related to $\mathscr{R}_T$ for $T \geqslant |A| - 1$, where $\mathscr{R}$ is derived from a class of "democratic preference functions."

3. *Transitivity Properties*

THEOREM A3. (a) *For any ordering function* $\mathscr{R}, (S, r) \in D$ *and* T *such that* $T \geqslant |S| - 1$ *then* $\mathscr{R}_T, \mathscr{R}'_T, \mathscr{R}''_T, \mathscr{R}'''_T$ *are, respectively, transitive, quasi-transitive, acyclical, quasi-transitive.*

(b) *For any ordering function (fully cyclical ordering function)* $\mathscr{R}$, $(S, r) \in D$ *and* $2 \leqslant T < |S| - 1$, $\mathscr{R}_T$ *and* $\mathscr{R}'_T$ ($\mathscr{R}''_T$ *and* $\mathscr{R}'''_T$) *are quasi-transitive (cyclical).*

Note that Theorem A3 also does not require purely binary preferences.

4. *Independence Properties*

THEOREM A4. (a) *For any ordering function* $\mathscr{R}$, $C(A, \dot{\mathscr{R}}_T)$ *and* $C(A, \dot{\mathscr{R}}'_T)$ *satisfy* III1(a,b) *for* $T \geqslant 2$ *and satisfy* III2 *for* $T \geqslant 2$ *and* $T = 2$ *respectively.*

(b) *If* $C(A, \mathscr{R})$ *satisfies* III1a *and* $\mathscr{R}$ *is transitive then* $\dot{\mathscr{R}}_T \geqslant \mathscr{R}$ *for all* $(S, r) \in D$ *and* T *that* $T \geqslant |S| - 1$.

Part (a) *states precisely the* III-*independence results shown in the length spectrum diagram. Part* (b) *implies that* III1(a) *cannot be achieved for any transitive ordering function without reducing its distinguishability to* less *than one of the unrestricted equalization methods. Thus,* III1(a) *cannot be achieved with transitivity if length restriction is also desired.*

5. *Non-Manipulability Properties of* $\mathscr{R}'_2$.

The literature on strategic manipulation of the social outcome via misrepresentation of preferences (see Gibbard [8], Satterthwaite [18], Gardendors [6]) has specified another class of negative results for social choice theory. This section presents a brief consideration of this relative to $\mathscr{R}'_2$, which is shown to guarantee certain non-manipulability properties for linear individual preferences, denoted (S, p). The notion of "relative majority" subsequently defined is partially motivated by Kalai *et al.* [13].

DEFINITION A2 (Relative Majority). (a) Given some ordering function $\mathscr{R}$, its purely direct preferences $\dot{\mathscr{R}}$ are implied by a *relative majority* $= RM$, if there is some $M \subset K$, such that if for any $\{x, y\} \subset X$ and $(S, P) \in D$ such that $xp_i y$ for all $i \in M$, $yp_j x$ for all $j \in K - M$; then xPy. $M \in RM$ means M is a relative majority.

(b) The following properties are assumed for relative majorities. (i) $K \in RM$. (ii) If $M \in RM$, $M \subset M'$, then $M' \in RM$. (iii) If $M \in RM$, then $K - M \notin RM$.

(c) A *Constant Sum* relative majority satisfies, for all $M \subset K$; either M or $K - M$ belongs to RM.

DEFINITION A3 (Clearly Beneficial Manipulation). (a) If $C(A, \mathscr{R}(S, p)) = \{x\}$, the social ordering $\mathscr{R}$ is *clearly beneficial* for $j \in K$ to manipulate if for some (S, p') which satisfies $(S, p_i) = (S, p'_i)$ for all $i \neq j$; $yp_j x$ for all $y \in C(A, \mathscr{R}(S, p')$.

Part b is taken from Gardenfors [6], for the case in which the choice set is not a singleton. It is also simplified to the case of linear orderings, $(S, p) \in D$.

(b) $C(A, \mathscr{R}(S, p))$ is clearly beneficial for $j \in K$ to manipulate if for some (S, p') such that $(S, p_i) = (S, p'_i)$ for all $i \neq j$;

$C(A, \mathscr{R}(S, p')) >_j C(A, \mathscr{R}(S, p))$, where the partial orderings, $>_j$, between elements of β is defined: $A >_j B$, where $A, B \in \beta$, if and only if: (i) $A \subset B$ and for all $x \in A$, $y \in B - A$, $xp_j y$. (ii) $B \subset A$, and for all $x \in A - B$, $y \in B$, $xp_j y$. (iii) Neither $A \subset B$, $B \subset A$, nor $A = B$, and for all $x \in A - B$, $y \in B - A$, $xp_j y$.

THEOREM A5 (Non-Manipulability). *Let (S, r) be restricted to sets of linear orderings over $S \in \beta$, denoted (S, p). Assume an ordering function's purely binary preference $\dot{\mathscr{R}}$ are derived from constant sum realtive majorities defined in* DA2c, *and assume $C(A, \dot{\mathscr{R}}'_2(A, p)) = \{x\}$. Then $C(A, \dot{\mathscr{R}}'_2(A, p))$ is not clearly beneficial to manipulate as defined in* DA3a.

Remarks. (1) The theorem implies that a traditional majority or Condorcet winner, when it occurs, will be chosen and "stable" with respect to strategic manipulation as defined in DA3a.

(2) In the presence of cycles, $\dot{\mathscr{R}}'_2$ may not guarantee stability as defined in DA3b. It can be shown, for $T \geqslant |S| - 1$, that $C(A, \dot{\mathscr{R}}_T)$, $C(A, \dot{\mathscr{R}}'_T)$ are non-manipulable from both specifications.

(3) It is interesting that the notion of discriminability or "decisiveness" also connects with the manipulability literature via Gardenfors' conjecture (see [7]) that anonymous, neutral, Condorcet, and non-manipulable (for linear individual preferences) methods must be very "undecisive." The weak ability to discriminate of $\mathscr{R}_T$ and $\mathscr{R}'_T$ for $T \geqslant |S| - 1$, thus, fits well with Gardenfors' conjecture. However, that $\dot{\mathscr{R}}'_2$ still retains relatively strong decisive properties (even though only DA3a is guaranteed) can be regarded, to a limited degree, as a counter example to this conclusion.

APPENDIX B. PROOFS

The following notation is introduced to facilitate proofs: $x\dot{\mathscr{R}}^m(S, r)y$ if and only if $\overrightarrow{x\dot{y}}_m \subset S$; $x\dot{\hat{\mathscr{R}}}^m(S, r)y$ if and only if $x\dot{\hat{y}}_m \subset S$; and $x\dot{P}^m(S, r)y$ if and only if $x\hat{y}_m \subset S$, where the notation $[\overrightarrow{x\dot{y}}_m \subset S,\ x\dot{\hat{y}}_m \subset S,\ x\dot{\hat{y}}_m \subset S]$ is defined just prior to Definition 3.1 of the main text. Also, the following Lemmas are used in the proofs below. The proofs of Lemmas 1–4 are straightforward and omitted.

LEMMA 1.

$$\textit{If}\begin{bmatrix} \text{not } x\dot{\mathscr{R}}^m(S, r)\ y \\ \text{not } x\dot{P}^m(S, r)\ y \end{bmatrix} m \geqslant 1, \textit{ then for all } S' \subset S, \begin{bmatrix} \text{not } x\mathscr{R}^m(S', r)\ y \\ \text{not } x\dot{P}^m(S', r)\ y \end{bmatrix}.$$

LEMMA 2.

$$\textit{If}\begin{bmatrix} x\dot{\mathscr{R}}^m\ y \\ x\dot{\hat{\mathscr{R}}}^m\ y \end{bmatrix}, \textit{ where } x, y \in A, \begin{bmatrix} y \in C(A, \dot{\mathscr{R}}_T) \\ y \in C(A, \dot{\mathscr{R}}'_T) \end{bmatrix}$$

$$\textit{and } 1 \leqslant m \leqslant T, \textit{ then } \begin{bmatrix} x \in C^*(A, \dot{\mathscr{R}}_T) \\ x \in C^*(A, \dot{\mathscr{R}}'_T) \end{bmatrix}.$$

LEMMA 3.

$$\textit{If}\begin{bmatrix} x\dot{P}_T(S, r)\ y \\ x\dot{P}'_T(S, r)\ y \end{bmatrix} T \geqslant 1, \textit{ then for all } S' \subset S; \begin{bmatrix} x\dot{P}_T(S', r)\ y \\ x\dot{P}'_T(S', r)\ y \end{bmatrix}.$$

LEMMA 4. *If* A *is finite,* $\mathscr{R}$ *is quasi-transitive over* A, *and* $x \notin A - C(A, \mathscr{R})$; *then there exists some* $y \in C(A, \mathscr{R})$ *such that* yPx.

LEMMA 5.

$$\textit{If}\begin{bmatrix} x\mathscr{R}^m y,\ y\mathscr{R}^n z \\ xP^m y,\ yP^n z \end{bmatrix} \textit{then for some } 1 \leqslant t \leqslant |S| - 1, \begin{bmatrix} x\mathscr{R}^t z \\ xP^t z \end{bmatrix}.$$

The proof for $\mathscr{R}^m$ is given, with that for P^m implied by an analagous argument. The definition of $\mathscr{R}^m$ requires the sequences of options which it connects to be distinct. Thus, an option "a" can occur at most once in both sequences which connect x to y and y to z in $x\mathscr{R}^m y$, and $y\mathscr{R}^n z$, respectively. If such an "a" occurs, then form $x\mathscr{R}^s a$, $a\mathscr{R}^t z$; and repeat this process recursively until for some a', $x\mathscr{R}^u a'$, $a'\mathscr{R}^v z$, the sequences connecting x to a' and a' to z are disjoint. Then, the union of these two sequences, which together connect x to z, will be distinct. This implies $x\mathscr{R}^{u+v}z$. Now since the number of pairwise preferences between any two options is one less than the number of options in the sequence (including x and z); then (since the sequence is distinct) $u+v$ can be no larger than $|S|-1$.

THEOREM A1. *$\mathscr{R}'''_T > \mathscr{R}''_T > \mathscr{R}'_T > \mathscr{R}_T$ for all $T \geqslant 2$ and $\mathscr{R}$ fully cyclical. It is easily shown that $\mathscr{R}'''_T \geqslant \mathscr{R}''_T \geqslant \mathscr{R}'_T \geqslant \mathscr{R}_T$. Then it remains to prove strictly greater distinguishability; since $\mathscr{R}$ is fully cyclical the following cases can be constructed.*

Assume xPy, yPz, xIz. Then we have $xP'''_T y$, $yP'''_T z$, $xP'''_T z$ for $T \geqslant 2$; $xP''_T y$, $yP''_T z$, $xI''_T z$, for $T \geqslant 2$; $xI'_T y$, $yI'_T z$, $xI'_T z$ for $T \geqslant 2$; $xI_T y$, $yI_T z$, $xI_T z$ for $T \geqslant 2$. This implies $\mathscr{R}'''_T > \mathscr{R}''_T$, $\mathscr{R}_T''' > \mathscr{R}'_T$, $\mathscr{R}'''_T > \mathscr{R}_T$, $\mathscr{R}''_T > \mathscr{R}'_T$, $\mathscr{R}''_T > \mathscr{R}_T$ for all $T \geqslant 2$. Now assume xIy, yIz, xPz. Then we have, $xI'_T y$, $yI'_T z$, $xP'_T z$ for $T \geqslant 2$: $xI_T y$, $yI_T z$, $XI_T z$ for $T \geqslant 2$. This proves $\mathscr{R}'_T > \mathscr{R}_T$ for all $T \geqslant 2$.

THEOREM A2. *For all $(A, r) \in D$ and T such that $T \geqslant |A|-1$; $C_{\mathrm{Sch}}(A, \mathscr{R}(A, r)) = C(A, \mathscr{R}'''_T(A, r))$. Since the theorem is essentially equivalent to Deb's theorem (see [5]), the proof is here omitted. However, it is noted the proof would need to be modified to explicitly include the unrestricted length hypothesis $T \geqslant |A|-1$, which Deb's proof only implicitly assumes.*

THEOREM A3. (1) *It is first shown $\mathscr{R}_T$ is transitive for $T \geqslant |S|-1$. Assume $x\mathscr{R}_T y$ and $y\mathscr{R}_T z$, which implies for some $1 \leqslant m$, $n \leqslant T$, $x\mathscr{R}^m y$, $y\mathscr{R}^n z$.* By Lemma 5 for some $1 \leqslant t \leqslant |S|-1$, *$x\mathscr{R}^t z$. Thus, the definition of $\mathscr{R}_T$ immediately implies, $x\mathscr{R}_T z$.*

(2) *$\mathscr{R}'_T$ is quasi-transitive for all $T \geqslant 2$. Assume $xP'_T y$, $yP'_T z$. This implies: xPy and not $y\hat{\mathscr{R}}^m x$ for $2 \leqslant m \leqslant T$ and yPz and not $z\hat{\mathscr{R}}^m y$ for $2 \leqslant m \leqslant T$.*

Now, if $z\mathscr{R}x$ then $z\mathscr{R}x$, xPy implies $z\hat{\mathscr{R}}^2 y$, which implies $z\mathscr{R}'_2 y$; thus contradicting the hypothesis. Thus, $z\mathscr{R}x$ is false, and xPz. Alternatively, assume $z\hat{\mathscr{R}}^m x$ for $2 \leqslant m \leqslant T$. Then for some $\{a_0, a_1, \ldots, a_{m-1}\} \subset S$. $a_0 = z$; $z\mathscr{R}a_1, \ldots, a_{m-1}\mathscr{R}x$, where at least one of the pairwise preferences is strict.

There are two cases. Either $a_{t-1}Pa_t$ *for* $1 \leqslant t \leqslant m-1$ *or* $a_{m-1}Px$. *Assume the former case. If* $a_t\dot{\mathscr{R}}y$ *then* $z\hat{\dot{\mathscr{R}}}^{t+1}y$, *where* $2 \leqslant t+1 \leqslant m$. *Thus,* $x\dot{\mathscr{R}}'_T y$ *which contradicts the hypothesis. If* yPa_t *then* $y\hat{\dot{\mathscr{R}}}^{m-t+1}x$, *where* $2 \leqslant m-t+1 \leqslant m$, *which implies* $y\dot{\mathscr{R}}'_T x$ *which contradicts the hypothesis. Hence, the remaining possibility is* $a_{m-1}Px$. *If* $y\dot{\mathscr{R}}a_{m-1}$ *then* $y\hat{\dot{\mathscr{R}}}^2 x$ *which implies the contradiction* $y\dot{\mathscr{R}}'_T x$. *If* $a_{m-1}Py$ *then* $z\hat{\dot{\mathscr{R}}}^m y$, *which also implies the contradiction* $z\dot{\mathscr{R}}'_T y$. *Hence* $z\hat{\dot{\mathscr{R}}}^m x$ *must be false, which together with* xPz, *imply* $xP'_T z$.

(3) $\dot{\mathscr{R}}''_T$ *is acyclical for* $T \geqslant |S|-1$. *Assume* $a_0P''_Ta_1,\ldots,a_{m-1}P''_Ta_m$ *where* $a_0,\ldots,a_m$ *are distinct and* $2 \leqslant m \leqslant |S|-1$. *Now this immediately implies by the definition of* $\dot{\mathscr{R}}''_T$, $a_0Pa_1,\ldots,a_{m-1}Pa_m$, *which implies* $a_0P^m a_m$. *Since* $m \leqslant |S|-1 \leqslant T$, *then* $a_0\dot{\mathscr{R}}''_T a_m$.

(4) $\dot{\mathscr{R}}'''_T$ *is quasi-transitive for* $T \geqslant |S|-1$. *Assume* xP'''_Ty, yP'''_Tz. *This implies:* $xP^m y$ *for* $1 \leqslant m \leqslant T$ *and not* $yP^n x$ *for all* $1 \leqslant n \leqslant T$, $yP^s z$ *for* $1 \leqslant s \leqslant T$ *and not* zPy^t *for all* $1 \leqslant t \leqslant T$. *By Lemma 5,* $xP^u z$, *for* $1 \leqslant u \leqslant T$. *Now if* $zP^v x$ *is also true then Lemma 5 implies,* $zP^q y$ *for some* $1 \leqslant q \leqslant T$, *which contradicts the hypothesis, since* $z\dot{\mathscr{R}}_T'''y$ is then implied. Hence we have $xP^u z$ *for* $1 \leqslant u \leqslant T$ *and not* $zP^v x$ *for all* $1 \leqslant v \leqslant T$ *which implies* $xP'''_T z$.

(5) *If* $\dot{\mathscr{R}}$ *is fully cyclical for* S *and* $T < |S|-1$, *then* $\dot{\mathscr{R}}''_T$ *and* $\dot{\mathscr{R}}'''_T$ *are cyclical.*

Since $\dot{\mathscr{R}}$ *is fully cyclical the following case can be constructed.*

For $\{a_0, a_1,\ldots,a_{m-1}\} = S$, *where* $m = |S|$: $a_0P_1,\ldots,a_{m-2}Pa_{m-1}, a_{m-1}Pa_0$ *and for all* a_i, a_j, *such that* $i, j \in \{0, 1,\ldots, m-1\}$ *and* $j \neq i+1$, a_iIa_j. *Now for* $T < |S|-1$ *this case implies for all* a_i, a_{i+1}, $i = 0, 1,\ldots, m-1$; a_iPa_{i+1} *and not* $a_{i+1}P^n a_i$ *for all* $1 \leqslant n \leqslant |S|-1$. *Thus:* $a_0P''_Ta_1,\ldots, a_{m-2}P''_Ta_{m-1}, a_{m-1}P''_Ta_0, a_0P'''_Ta_1,\ldots,a_{m-2}P'''_Ta_{m-1}, a_{m-1}P'''_Ta_0$.

THEOREM A4. (1) $C(A,\dot{\mathscr{R}}_T)$, $C(A,\dot{\mathscr{R}}'_T)$ *satisfy* III 1(a,b) *for all* $T \geqslant 2$. *Proof for* III 1(a). *In both cases the proof for* $\dot{\mathscr{R}}'_T$ *is given with similar argument applicable to* $\dot{\mathscr{R}}_T$. *Let* $x \in A - C^*(A, \dot{\mathscr{R}}'_T(S,r))$ *which implies for all* $y \in C(A, \dot{\mathscr{R}}'_T(S,r))$, $y\dot{P}'_T(S,r)\,x$. *Since by hypothesis* $C(A, \dot{\mathscr{R}}'_T(S,r)) - B \neq \varnothing$, *then there exists* $z \in C(A, \dot{\mathscr{R}}'_T(S,r))$ *such that* $z \in A - B$. *By Lemma* 3, $z\dot{P}'_T(S-B,r)\,x$, *which implies* $C(A-B, \dot{\mathscr{R}}'_T(S-B,r)) \subset C(A-B, \dot{\mathscr{R}}'_T(S-B,r))$.

Proof of III 1(b). *The hypothesis implies* $C(A, \dot{\mathscr{R}}'_T(S,r)) \subset S - B$. *Let* $x \in A - C(A, \dot{\mathscr{R}}'_T(S,r))$, *which implies by Lemma 4 and Theorem* A3 *(quasi-transitivity of* $\dot{\mathscr{R}}'_T$ *for* $T \geqslant 2$); *there exists* $y \in C(A, \dot{\mathscr{R}}'_T(S,r))$ *such that* $y\dot{P}'_T(S,r)\,x$. *Since* $y \in A - B$, *then Lemma 3 implies* $y\dot{P}'_T(S-B,r)\,x$. *Thus,* $x \notin C(A-B, \dot{\mathscr{R}}'_T(S-B,r))$ *which implies* $C(A-B, \dot{\mathscr{R}}'_T(S-B,r)) \subset C(A, \dot{\mathscr{R}}'_T(S,r))$. *Now assume* $x \in C(A, \dot{\mathscr{R}}'_T(S,r))$, *which implies either* $x\dot{\mathscr{R}}y$, *or* $x\hat{\dot{\mathscr{R}}}^m(S,r)\,y$, $2 \leqslant m \leqslant T$, *for all* $y \in A$. *If* $x\dot{\mathscr{R}}(S,r)\,y$ *then the definition* *of* $\dot{\mathscr{R}}$ *implies* $x\dot{\mathscr{R}}(S-B, r)\,y$, *which implies* $x\dot{\mathscr{R}}'_T(S-B, r)\,y$. *Alternatively assume* $y\dot{P}(S, r)\,x$ *and note that* $x\hat{\dot{\mathscr{R}}}^m(S, r)\,y$ *implies for some* $\{a_1, \dots, a_{m-1}\} \subset S$ *that* $x\dot{\mathscr{R}}a_1, \dots, a_{m-1}\dot{\mathscr{R}}y$. *Thus, for all* $a_i \in \{a_1, \dots, a_{m-1}\}$, $y\dot{P}x$ *implies* $a_i\hat{\dot{\mathscr{R}}}^m i(S, r)\,x$, *where* $m_i = m - i + 1$ *where* $2 \leqslant m_i \leqslant m \leqslant T$. *Thus Lemma 2 implies* $a_i \in C^*(A, \dot{\mathscr{R}}_T(S, r))$ *for all* $i = 1, \dots, m-1$. *Since the hypothesis implies* $C^*(A, \dot{\mathscr{R}}_T(S, r)) \subset S - B$, *then,* $y \in A - B$, *we have* $x\hat{\dot{\mathscr{R}}}^m(S-B, r)\,y$, *where* $m \leqslant T$. *Therefore,* $x\dot{\mathscr{R}}'_T(S-B, r)\,y$, *in both cases* $x\dot{\mathscr{R}}y$ *and* $y\dot{P}x$, *for all* $y \in A - B$. *Thus* $C(A, \dot{\mathscr{R}}'_T(S, r)) \subset C(A - B, \dot{\mathscr{R}}'_T(S - B, r))$

(2) $C(A, \dot{\mathscr{R}}_2)$, $C(A, \dot{\mathscr{R}}'_2)$ *satisfy III 2. The proof given also applies to* $\dot{\mathscr{R}}_2$. *If* $x \in A - C(A, \dot{\mathscr{R}}'_2(S, r))$ *then Lemma 4 implies* $y\dot{P}'_2(S, r)\,x$ *for some* $y \in C(A, \dot{\mathscr{R}}'_2(S, r))$. *Thus,* $y\dot{P}(S, r)\,x$ *and not* $x\hat{\dot{\mathscr{R}}}^2(S, r)\,y$. $ext[C^*(A, \dot{\mathscr{R}}'_2(S, r)), r, r']$ *implies* $(\{x, y\}, r) = (\{x, y\}, r')$, *which implies* $y\dot{P}(S, r')\,x$. *Now, in order to have* $x\dot{\mathscr{R}}'_2(S, r')\,y$, *this means:* $x\dot{P}(S, r')\,a$, $a\dot{\mathscr{R}}(S, r')\,y$ *or* $x\dot{\mathscr{R}}(S, r')\,a$, $a\dot{P}(S, r')\,y$. *In either case,* $a\dot{\mathscr{R}}(S, r')\,y$ *is required. Thus, since* $ext[C^*(A, \dot{\mathscr{R}}'_2(S, r)), r, r']$ *implies* $(\{a, y\}, r) = (\{a, y\}, r')$; *then not* $a\dot{\mathscr{R}}(S, r)\,y$ *implies not* $a\dot{\mathscr{R}}(S, r')\,y$. *Hence,* $a\dot{\mathscr{R}}(S, r')\,y$, $a\dot{P}(S, r')\,y$ *imply* $a\dot{\mathscr{R}}(S, r)\,y$, $a\dot{P}(S, r)\,y$, *respectively. This implies* $a\dot{\mathscr{R}}_2(S, r)\,y$, *which means* $aI'_2(S, r)\,y$; *thus implying* $a \in C^*(A, \dot{\mathscr{R}}'_2(S, r)$ *by Lemma 2. This implies by* ext$[C^*(A, \dot{\mathscr{R}}'_2(S, r)), r, r']$ *that* $(\{x, a\}, r) = (\{x, a\}, r')$. *Thus,* $x\dot{P}(S, r')\,a$, $x\dot{\mathscr{R}}(S, r')\,a$ *imply* $x\dot{P}(S, r)\,a$, $x\dot{\mathscr{R}}(S, r)\,a$, *respectively. Therefore:*

$$[x\dot{P}(S, r')\,a, a\dot{\mathscr{R}}(S, r')\,y] \text{ implies } [x\dot{P}(S, r)\,a, a\dot{\mathscr{R}}(S, r)\,y], \tag{1}$$

$$[\dot{\mathscr{R}}(S, r')\,a, a\dot{P}(S, r')\,y] \text{ implies } [x\dot{\mathscr{R}}(S, r)\,a, a\dot{P}(S, r)\,y]. \tag{2}$$

From Eqs. (1) and (2), $x\hat{\dot{\mathscr{R}}}^2(S, r')\,y$ *implies* $x\hat{\dot{\mathscr{R}}}^2(S, r)\,y$, *which implies* $x\dot{\mathscr{R}}_2(S, r)\,y$. *But this contradicts* $y\dot{P}'_2(S, r)\,x$ *previously implied. Thus, we have* $y\dot{P}(S, r')\,x$ *and not* $x\hat{\dot{\mathscr{R}}}^2(S, r')\,y$: *which implies* $y\dot{P}'_2(S, r')\,x$. We have thus shown $C(A, \dot{\mathscr{R}}'_2(S, r')) \subset C(A, \dot{\mathscr{R}}'_2(S, r))$. *The same argument applied to* $x \in A - C(A, \dot{\mathscr{R}}'_2(S, r'))$ *implies* $C(A, \dot{\mathscr{R}}'_2(S, r)) \subset C(A, \dot{\mathscr{R}}'_2(S, r'))$ *by noting that the first result implies* ext$[C^*(A, \dot{\mathscr{R}}'_2(S, r')), r, r']$.

(3) $C(A, \dot{\mathscr{R}}_T(S, r))$ *satisfies III2 for all* $T \geqslant 2$.

The above argument can be extended in recursive fashion for $\dot{\mathscr{R}}_T$, $T \geqslant 2$ by using Lemma 2. The details of the proof are omitted.

THEOREM 4.1a. $C(A, \mathscr{R}(S, r))$ *satisfies IIA if and only if* $\mathscr{R} = \dot{\mathscr{R}}$. *The proof follows directly from the definitions of* $C(A, \mathscr{R})$, $\dot{\mathscr{R}}$, *and IIA; and is omitted.*

THEOREM 4.1b. (1) *If* $C(A, \mathscr{R}(S, r))$ *satisfies III* 1(a), *then* $\mathscr{R}$ *and* $\dot{\mathscr{R}}$ *are companions.*

Assume $xP(S,r)y$ *which implies* $y \notin C^*(\{x,y\}, \mathscr{R}(S,r))$. *Thus, by* III 1(a), *and since* $C(\{x,y\}, \mathscr{R}(S,r))$ *has only one element,* x; $C(\{x,y\}, \mathscr{R}(\{x,y\},r)) = \{x\}$. *This implies* $xP(\{x,y\},r)y$. *Thus* xPy *obviously implies not* $y\dot{P}x$; *and if* $x\dot{P}y$, *then not* yPx *since otherwise* $y\dot{P}x$ *is implied.*

THEOREM 5.1. *If* $\dot{\mathscr{R}}$ *of* $\mathscr{R}$ *satisfies, neutrality, strong Pareto Principle, and positive association, then* $\dot{\mathscr{R}}'_2$, $\dot{\mathscr{R}}''_2$, $\dot{\mathscr{R}}'''_2$ also guarantee the strong Pareto Principle.

The proof is given for part (2) of the Pareto principle, with similar argument applying to part (1). Assume for some $(S,r) \in D$ and some x, $y \in S$ that xr_iy for all i and xp_iy for some i. It is shown, not $x\dot{P}'_2y$ leads to a contradiction. Since $\dot{P}$ by hypothesis satisfies the strong Pareto principle, then $x\dot{P}y$. Thus, not $x\dot{P}'_2y$ implies either (a) $y\dot{\mathscr{R}}z$ and $z\dot{P}x$, or (b) $y\dot{P}z$, $z\dot{\mathscr{R}}x$ for some $z \in S$ which is distinct from x and y. Now, suppose (a) holds. Since for all i, xr_iy then by transitivity of each r_i, zp_ix implies zp_iy and zr_ix implies zr_iy, for all i. Thus, by neutrality and positive association $z\dot{P}x$ implies $z\dot{P}y$, which contradicts $y\dot{\mathscr{R}}z$ of (a). Similarly, (b) also leads to contradiction.

Since $\dot{\mathscr{R}}''_2, \dot{\mathscr{R}}'''_2$ are both at least as discriminating as $\dot{\mathscr{R}}'_2$, the above argument also applies

THEOREM 5.2. *Parts* (a), (b), (c), (d) *are already implied by Theorems* A1–A4, 5.1. *Part*(e) *is easily proven from the definitions of conditions* 4 *and* 5 *of Arrow* [1].

THEOREM A5. *For linear individual preferences,* (S,p); $C(A, \dot{\mathscr{R}}'_2(A,p)) = \{x\}$, *and* $\dot{\mathscr{R}}$ *derived from constant-sum relative majorities*; $C(A, \dot{\mathscr{R}}'_2(A,p))$ *is not clearly beneficial to manipulate as defined in* D19a. *The hypothesis imply that for all* (S,p) *and all distinct* x, $y \in A$; $x\dot{P}y$ *or* $y\dot{P}x$. *The following Lemma is first proved.*

LEMMA 6. 6. *If* $C(A, \dot{\mathscr{R}}'_2(A,p)) = \{a\}$ *then for all* $b \in A$ *distinct from* a, $a\dot{P}b$. *For any such* $b \in A$, $b \neq a$, *since* $\dot{\mathscr{R}}'_2$ *is quasi-transitive by Theorem* A3, *then Lemma* 4 *implies* $a\dot{P}'_2b$, *which implies by the definition of* $\dot{\mathscr{R}}'_2$, $a\dot{P}b$.

To prove Theorem A5, assume the hypothesis and assume there exists $x \in C(A, \dot{P}'_2(S,r^0))$, such that xp_ja for individual $j \in K$. To simplify notation let $P = \dot{P}'_2(S,r)$, $P^0 = \dot{P}'_2(S,r^0)$, $\dot{P} = \dot{P}(S,r)$ $\dot{P}^0 = \dot{P}(S,r^0)$. By Lemma 6, $a\dot{P}x$ and since xp_ja, $a\dot{P}^0x$. Since $x \in C(A,P^0)$, then $a\dot{P}^0x$ implies there exists $y \in A$ such that $x\dot{P}^0y$, $y\dot{P}^0a$. Since by Lemma 6, $a\dot{P}y$ but $y\dot{P}^0a$, individual j must have changed his ordering relative to $\{a,y\}$, which implies his initial preference must have been, ap_jy. Thus, if $y \in C(A, P^0)$, then clearly beneficial manipulation is violated. Thus, $y \notin C(A, P^0)$ which implies there exists $z \in A$ such that $z\dot{P}^0y$ and not $y\hat{\mathscr{R}}^0z$, where $\hat{\mathscr{R}}^0 = \hat{\mathscr{R}}^2(S, r^0)$. Since $y\dot{P}^0a$ and $a\dot{P}^0x$, then not $y\hat{\mathscr{R}}^0z$ implies $z \neq x; y\dot{P}^0a$ also implies $z \neq a$, and obviously $z \neq y$, since $z\dot{P}^0y$. Thus, z is distinct from a, x, y. Since $y\dot{P}^0a$, then not $y\hat{\mathscr{R}}^0z$ implies not $a\dot{P}^0z$ which implies $z\dot{P}^0a$. But since, by Lemma 6, $a\dot{P}z$ then j must have changed his ordering, which implies his initial preference must have been ap_jz. Thus, if $z \in C(A, P^0)$, clearly beneficial manipulation is again violated. Thus, $z \notin C(A, P^0)$ which implies for some $z' \in A$, $z'\dot{P}^0z$ and not $z\hat{\mathscr{R}}^0z'$. As above it is readily shown that z' is distinct from a, x, y, z and j's initial preference must be ap_jz'. Thus, for some $z'' \in A$, $z''\dot{P}^0z'$, and not $z'\hat{P}^0z'$, and not $z'\hat{P}^0z''$; which sequence of distinct options $(z, z'z'',\ldots,)$ must proceed indefinitely. But this contradicts the finiteness of A; and therefore the assumption $x \in C(A, P^0)$, which started it, must be false.

ACKNOWLEDGMENTS

Comments made by an associate editor were helpful. All inadequacies rest with the author.

REFERENCES

1. K. J. Arrow, "Social Choice and Individual Values," 2nd Ed., Yale Univ. Press, New Haven, Conn., 1963.
2. K. J. Arrow, Rational choice functions and orderings, *Econometrica* **26** (1959).
3. R. Batra and P. K. Pattanaik, On some suggestions for having non-binary social choice functions, *Theory and Decision* **3** (1972), 1–11.
4. D. E. Cambell, Democratic preference functions, *J. Econ. Theory* **12** (1976), 259–272.
5. R. Deb, On Schwartz's rule, *J. Econ. Theory* **16** (Oct. 1977].
6. P. Gardenfors, Positionalist voting functions, *J. Econ. Theory* **13** (1976), 217–118.
7. P. Gardenfors, Manipulation of voting functions, *J. Econ. Theory* **13** (1976), 117–118.
8. A. Gibbard, Manipulation of voting schemes: A general result, *Econometrica* **41** (1973), 558–574.
9. B. Hansson, The independence condition in the theory of social choice, *Theory and Decision A* (1973), 25–49.
10. R. A. Heiner, "The Meaning of IIA," Department of Economics Research Paper, Brigham Young University, Provo, Utah, November 1978.
11. R. A. Heiner and D. J. Packard, Ranking functions and independence conditions, *J. Econ. Theory* **16** (Oct. 1977).
12. R. A. Heiner, "Cycles and Strength Priority Preferences," Department of Economics Research Paper, June 1979, Brigham Young University, Provo, Utah, *J. Econ. Theory*, submitted.
13. R. Kalai, E. Pazner, and D. Schmeidler, Collective choice correspondences as admissible outcomes of social bargaining processes, *Econometrica* **44** (1976), 233–240.
14. G. Kramer, On a class of equilibrium conditions for majority rule, *Econometrica* **41** (1973).

15. G. Kramer, A note on single peakedness, *Internat. Econ. Rev.* **17** (1976), 498–502.
16. R. D. McKelvey, Intransitivities in multi-dimensional voting models and some implications for agenda control, *J. Econ. Theory* **12** (1976), 472–482.
17. V. S. Ramachandran, Liberalism, non-binary choice and the Pareto principle, *Theory and Decision* **3** (1972), 49–54.
18. M. Sattzerthwaite, Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *J. Econ. Theory* **10** (1975), 187–217.
19. T. Schwartz, Rationality and the myth of the maximum, *Nous* (May 1972).
20. T. Schwartz, Choice functions, rationality conditions, and variations on the weak axiom of revealed preference, *J. Econ. Theory* **16** (1977).
21. A. K. Sen, Choice functions and revealed preference, *Ref. Econ. Statist.* (1973), 307–317.
22. A. K. Sen, "Collective Choice and Social Welfare," Holden-Day, San Francisco, 1971.
23. P. Wagstaff, Proof of a conjecture in social choice, *Internat. Econ. Rev.* **17** (1976), 491–497

JOURNAL OF ECONOMIC THEORY 25, 131–143 (1981)

Notes, Comments, and Letters to the Editor

Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement

FRANK A. COWELL

London School of Economics, London WC2A, 2AE, United Kingdom

AND

KIYOSHI KUGA

Osaka University, Suita, Osaka, 565, Japan

Received October 3, 1977; revised May 23, 1980

1. INTRODUCTION

In view of the variety of specifications to be found in the literature it is clear that there is no unequivocal rule for the selection of an "appropriate" inequality measure. Whether such measures are derived from social welfare theory or from other considerations with intuitive appeal one usually finds that the reasonable properties which are proposed in each case are insufficient for a unique ranking of income distributions. In addition practical inequality measures often take the form of an additive function of incomes or income shares—the Gini coefficient being an obvious exception. These observations provide the motivation for the two main questions of this paper.

The first question is—Why use additive inequality measures? We answer this in Section 2 by providing a simple axiomatic treatment of inequality measurement, the most crucial step being the assumption that inequality can be decomposed by population subgroup. The second question is—What more is required to provide a specific operational inequality measure? In Section 3 two extra axioms relating to the effects of additions to the population yield the "generalized entropy" class which subsumes the inequality measures proposed by Atkinson [1] and Theil [11] as special cases. Finally in Section 4 we show that a further axiom of maximum sensitivity leaves us with Theil's measure alone.

0022-0531/81/040131-13$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

2. An Additivity Theorem

As a preliminary step we shall provide a set of sufficient conditions under which an inequality measure takes an additive form. By an "inequality measure" we mean a real valued function defined on the space of "income shares" for any given population number, and having the following properties:

Axiom 1 (Scale Independence). *The inequality measure* $M(s; n)$ *is a function which may depend on the population size* n *and which is defined on the space of income shares, where* $s \in S^n = \{s = (s_1, \ldots, s_n) \geqslant 0 \mid \sum_{i=1}^{n} s_i = 1\}$.

Axiom 2 (Anonymity). $M(s; n) = M(Ps; n)$, $s \in S^n$, *where* P *is any arbitrary permutation matrix of size* $n \times n$.

These two axioms would command wide acceptance, the first stating that inequality is unaltered by proportional changes in all incomes,[1] and hence is independent of the units of measurement of income, and the second that inequality is independent of which labels are assigned to which income shares in the population. The next axiom incorporates a principle first introduced by Pigou [8] and Dalton [3]—a transfer from a rich man to a poor man, that is not so large as to reverse their positions, should definitely reduce inequality. In order to formulate the principle in a limiting form that is convenient for the differential calculus, we shall define

$$T_{ji}(s) \equiv \lim_{\varepsilon \to 0} \{M(s_1, \ldots, s_i + \varepsilon, \ldots, s_j - \varepsilon, \ldots, s_n; n) - M(s; n)\}/\varepsilon$$

and assume

Axiom 3 (Principle of Transfers). $T_{ji}(s) < 0$ *for* $s_i < s_j$ *where it exists.*

Regarding the transfer operation there is one other aspect that requires explanation. Since we are concerned with the definition of a measure on income shares, an increase in one man's income and a concurrent decrease in the income of some other one cannot be dealt with by partial differentiation. Formally stated, since s belongs to the unit simplex S^n on which $M(s; n)$ is defined, partial differentiation is not operational on S^n.[2] To avoid this problem, define

---

[1] Some authors have reservations about this, however—see Kolm [6].

[2] It might be thought that this problem can be avoided by defining $M(s; n)$ on R_+^n, the entire non-negative orthant of Euclidean n-space, imposing zero-order homogeneity as an additional axiom, and using conventional separability analysis. This is unsatisfactory for two reasons. Firstly the conventional separability arguments in Goldman and Uzawa [5] and

$$z = (z_1, \ldots, z_{n-1}) \in Z = \left\{ z \geqslant 0 \;\middle|\; \sum_{i=1}^{n-1} z_i \leqslant 1 \right\},$$

$$z_i = s_i \ (i = 1, \ldots, n-1), \qquad s_n = 1 - \sum_{i=1}^{n-1} z_i,$$

and write

$$M(s; n) = M\left(z_1, \ldots, z_{n-1}, 1 - \sum_{i=1}^{n-1} z_i ; n\right) = m(z; n).$$

Under this representation we have[3]

$$T_{ni}(s) = \frac{\partial}{\partial z_i} m(z; n) \equiv m_i(z; n); \qquad i = 1, \ldots, n-1. \tag{1}$$

We are now ready to introduce

AXIOM 4 (Smoothness). *$m(z; n)$ has* continuous *partial derivatives on* $\hat{Z} = \{z > 0 \mid \sum_{i=1}^{n-1} z < 1\}$ up to the third order *with the commutative property.*

In an application of inequality measurement it is desirable not only to compare overall levels of inequality but also to examine the separate contributions to differences in inequality between the two distributions. To make this possible it is useful to be able to decompose the inequality index by population subgroups.[4] Accordingly, consider a decomposition of the population $N = \{1, \ldots, n\}$ into two subpopulations N_1 and N_2 in such a way that $N_1 \cap N_2 = \emptyset$ and $N_1 \cup N_2 = N$. Let n_1 and n_2 be the numbers in the two groups, let s^1 and s^2 be the subvectors of s, the components of which correspond to the shares in total income of the members of groups N_1 and N_2, respectively, and let $\sigma_1 \equiv \sum_{j \in N_1} s_j$, $\sigma_2 \equiv \sum_{j \in N_2} s_j$. We impose only a very weak decomposability requirement, namely:

---

elsewhere prove to be inconsistent with an a priori restriction of zero order homogeneity. Secondly, for some applications of the analysis extension to R^n_+ is meaningless. For the conventional definition of income we may certainly consider a total of arbitrary size so that a numerical function may be defined not only on the unit simplex but on a differentiable extension in R^n_+. However, in other examples the total is necessarily fixed—such as inequality of land tenure or the measurement of malapportionment in voting systems— and thus our general approach becomes necessary.

[3] More generally we are able to prove $T_{ji}(s) = m_i(z; n) - m_j(z; n)$; $i \neq n, j \neq n$.

[4] See Éltető and Frigyes [4, p. 384], Takayama [10].

AXIOM 5 (Decomposability). *For any arbitrary subdivision of the population such that* $n_1 \geqslant 3$ *and* $n_2 \geqslant 2$, *we may find functions* F, G, H, *possibly depending on* n, n_1, n_2, *respectively, such that*

$$M(s; n) = F(G(s^1; n_1), H(s^2; n_2); n_1, n_2, \sigma_1, \sigma_2). \tag{2}$$

This axiom extends Theil's idea [11] of writing total inequality of within-group and between-group inequality. The functions G and H may be simply transformations of $M(s^1; n_1)$ and $(s^2; n_2)$ the inequality measure defined on the appropriate spaces for the groups N_1 and N_2, although we do not actually need to make this restriction. Intergroup inequality is incorporated in the explicit dependence of F upon the group shares in total income σ_1 and σ_2.

We are now able to state

THEOREM 1. *Axioms 1–5 lead to an additive measure such that*

$$M(s; n) = f\left(\sum_{i=1}^{n} \theta(s_i; n); n\right), \tag{3}$$

where f is a monotonic transformation possibly depending on n, and $\theta(;)$ is a common scalar-to-scalar function also depending upon n

The proof is relegated to Appendix A.

3. THE GENERALIZED ENTROPY CLASS

The class of inequality measures to which we have been directed thus far is quite broad, namely, the class of thrice-differentiable, convex,[5] symmetric, additive functions from S^n to the real line. Different members of this class may have very different ordering properties when comparing points in S^n. Our purpose in this and the following section is to examine in what way the imposition of certain other plausible requirements for inequality measurement further restricts the class of derived inequality measures.

For our next axiom we need to consider the iso-inequality contours of the function $M(s; n)$—akin to the indifference curves of a conventional utility function. However, while the marginal rate of substitution on a conventional indifference curve is conventionally defined as ratio between Δs_i and Δs_j such that $M(s; n) = M(s + \Delta s; n)$, such a definition is invalid here because s, $s + \Delta s \in S^n$, and $\Delta s_i + \Delta s_j = 0$, so that we may have $M(s; n) \neq M(s + \Delta s; n)$ in view of Axiom 3. Hence we allow s_k, an arbitrary *third* income share, to vary such that $\Delta s_i + \Delta s_j + \Delta s_k = 0$, $M(s\,;n) = M(s + \Delta s\,;n)$, and define the *Marginal Rate of Substitution in Inequality* (MRSI) as

---

[5] Axiom 3 implies that $\theta''(x; n) \geqslant 0$ for all $x > 0$, with θ'' not identically zero in any non trivial interval.

$$-\frac{\Delta s_j}{\Delta s_i} = \frac{\{M[s + (\Delta s_i, \Delta s_j, \Delta s_k)\,;n] - M[s + (0, \Delta s_j, \Delta s_i + \Delta s_k)\,;n]\}/\Delta s_i}{\{M[s + (0, \Delta s_j, \Delta s_i + \Delta s_k);n] - M(s\,;n)\}/\Delta s_j}$$

$$= T_{ki}(s)/T_{kj}(s) \qquad \text{as} \quad \Delta s_i, \Delta s_j \to 0. \tag{4}$$

Hence the MRSI tells us the income-share decrease for j which exactly offsets a given income-share increase for i, the result possibly depending on n and which k is chosen to accommodate this change in shares. We now further restrict $M(s\,;n)$ by requiring that the MRSI remains unchanged when the population is increased by introducing an $(n+1)$th person with nil income. Though this requirement is less immediately obvious than Axioms one to five, its intuitive interpretation is straightforward—the inequality-compensating adjustments *within* the existing population should be independent of any newcomers that do not alter the existing income shares.[6] Since we must now denote explicit dependence on n we shall write $s(n) \in S^n$ instead of $s \in S^n$ and $T_{ji}(s(n); n)$ for $T_{ji}(s)$.

AXIOM 6 (Indifference curves). *The* MRSI *between* s_i *and* s_j *with accommodating change in* s_k *(whoever* k *is) remains unaffected when a newcomer enters the society with nil income.*

Such a situation can be represented by

$$\frac{T_{ki}(s(n+1); n+1)}{T_{kj}(s(n+1); n+1)} = \frac{T_{ki}(s(n); n)}{T_{kj}(s(n); n)} \qquad \text{with } s(n+1) = [s(n), 0] \in S^{n+1}.$$

By virtue of the axiom together with (4), we may put

$$\frac{\theta'(z_i\,;n) - \theta'(z_k\,;n)}{\theta'(z_j\,;n) - \theta'(z_k\,;n)} = \frac{\theta'(z_i\,;n+1) - \theta'(z_k\,;n+1)}{\theta'(z_j\,;n+1) - \theta'(z_k\,;n+1)}. \tag{5}$$

Differentiating (5) with respect to z_i and z_k, we have

$$\theta''(z_i\,;n+1)/\theta''(z_i\,;n) = \theta''(z_k\,;n+1)/\theta''(z_k\,;n). \tag{6}$$

---

[6] In fact the requirement is rather weak, for two reasons. Firstly, the class of $M(s; n)$ to which it leads has this independence property for newcomers with *any* nonnegative incomes whatsoever—we take the value $s_{n+1} = 0$ purely for convenience in the derivation. Secondly, reinspection of Axiom 5 will reveal that this is only a very slight modification to the case of variable population.

Since z_l and z_k are arbitrary, we may take (6) to be a constant, say $c(n+1; n)$, which depends only on n and $n+1$, and may rewrite it as

$$\theta''(x; n+1) = c(n+1; n)\,\theta''(x; n). \tag{7}$$

This relation may be used recursively beginning with, say $n = 4$, as follows:

$$\begin{aligned} &\theta''(x; 5) = c(5; 4)\,\theta''(x; 4), \\ &\theta''(x; \nu) = \gamma(\nu)\,\theta''(x; 4), \qquad \text{where} \quad \gamma(\nu) = \prod_{l=5}^{\nu} c(l; l-1). \end{aligned} \tag{8}$$

Integrating (8) twice we have

$$\theta(x; \nu) = \gamma(\nu)\,\tau(x) + d(\nu)x + \mu(\nu), \tag{9}$$

where $\tau(x) = \theta(x; 4)$ and $d(\nu)$ and $\mu(\nu)$ are constants of integration. Using (9), $\sum_{i=1}^{n} \theta(s_i; n)$ can be rewritten as

$$\begin{aligned} \sum_{i=1}^{n} \theta(s_i; n) &= \gamma(n) \sum_{i=1}^{n} \tau(s_i) + d(n) \sum_{i=1}^{n} s_i + n\mu(n) \\ &= \gamma(n) \sum_{i=1}^{n} \tau(s_i) + \eta(n); \qquad \eta(n) = d(n) + n\mu(n). \end{aligned}$$

We may thus rewrite (4) as

$$\begin{aligned} M(s; n) &= f\left(\gamma(n) \sum_{i=1}^{n} \tau(s_i) + \eta(n); n\right) \\ &= \bar{f}\left(\sum_{i=1}^{n} \tau(s_i; n)\right), \qquad \text{where} \quad \bar{f}(x; n) = f(\gamma(n)x + \eta(n); n). \end{aligned} \tag{10}$$

We are now nearly at the final stage of our argument. The axiom to be introduced then is the principle of population, which is as follows.

Let $s(n) = (s_1, \ldots, s_n) \in S^n$ and put

$$\underbrace{\left(\frac{s(n)}{a}, \ldots, \frac{s(n)}{a}\right)}_{a} \equiv s(n)[a] \in S^{an},$$

where a is an arbitrary positive integer. $s(n)[a]$ represents the incomes share vector when the original population $s(n)$ is replicated a times. We shall then postulate

Axiom 7 (Principle of Population). $M(s(n)[a]; na) = M(s(n); n)$ *for arbitrary positive integers* a.

Axiom 7 is a discrete version of the statement that the inequality measure depends upon the relative density rather than the absolute density.

Combining (10) with Axiom 7, we have

$$\mathcal{J}\left(\sum_{l=1}^{n} \tau(s_l); n\right) = \mathcal{J}\left(\alpha \sum_{l=1}^{n} \tau\left(\frac{s_l}{\alpha}\right); \alpha n\right). \tag{11}$$

Considering (11) in the Z-domain and differentiating it with respect to z_i and z_j, respectively, we have

$$\{\tau'(z_l) - \tau'(1 - \sum z_v)\}/\{\tau'(z_j) - \tau'(1 - \sum z_v)\}$$
$$= \left\{\tau'\left(\frac{z_l}{\alpha}\right) - \tau'\left(\frac{1 - \sum z_v}{\alpha}\right)\right\}\Big/\left\{\tau'\left(\frac{z_j}{\alpha}\right) - \tau'\left(\frac{1 - \sum z_v}{\alpha}\right)\right\}. \tag{12}$$

The same operation may be performed for k and j, and combining the resultant relation with (12), we have

$$\{\tau'(z_l) - \tau'(z_k)\}/\{\tau'(z_j) - \tau'(1 - \sum z_v)\}$$
$$= \left\{\tau'\left(\frac{z_l}{\alpha}\right) - \tau'\left(\frac{z_k}{\alpha}\right)\right\}\Big/\left\{\tau'\left(\frac{z_j}{\alpha}\right) - \tau'\left(\frac{1 - \sum z_v}{\alpha}\right)\right\}. \tag{13}$$

A relation analogous to (13) holds for l, m, and j, and may be combined with (13) itself to yield

$$\{\tau'(z_l) - \tau'(z_k)\}/\{\tau'(z_l) - \tau'(z_m)\}$$
$$= \left\{\tau'\left(\frac{z_l}{\alpha}\right) - \tau'\left(\frac{z_k}{\alpha}\right)\right\}\Big/\left\{\tau'\left(\frac{z_l}{\alpha}\right) - \tau'\left(\frac{z_m}{\alpha}\right)\right\}. \tag{14}$$

Differentiating (14) with respect to z_l, we shall obtain

$$\tau''(z_l)\Big/\left\{\frac{1}{\alpha}\tau''\left(\frac{z_l}{\alpha}\right)\right\} = \{\tau'(z_l) - \tau'(z_m)\}\Big/\left\{\tau'\left(\frac{z_l}{\alpha}\right) - \tau'\left(\frac{z_m}{\alpha}\right)\right\}.$$

The r.h.s. above depends upon z_l, z_m and α. But the first two of them are arbitrarily locatable, and it therefore follows that the l.h.s. is a function only of α, say $\alpha\tau''(x)/\tau''(x/\alpha) = \zeta(\alpha)$ and thus upon differentiation

$$G(x) = G(x/\alpha) \qquad \text{for} \quad x \in (0, 1) \text{ and arbitrary positive integer } \alpha, \tag{15}$$

where $G(x) = g((x)/\{xg'(x)\}$ with $g(x) = x\tau''(x)$.

We want to show first that $G(x)$ is constant over the set of rational numbers belonging to $(0, 1)$. Let m_1, m_2, m_3, m_4 be arbitrary positive integers such that m_1/m_2 and m_3/m_4 belong to $(0, 1)$. Then by (15) we have

$$G\left(\frac{m_1}{m_2}\right)=G\left(\frac{1}{m_1}\cdot\frac{m_1}{m_2}\right)=G\left(\frac{1}{m_2}\right)=G\left(\frac{1}{m_2}\cdot\frac{1}{m_4}\right)$$
$$=G\left(\frac{1}{m_2 m_3}\cdot\frac{m_3}{m_4}\right)=G\left(\frac{m_3}{m_4}\right).$$

Since $G(x)$ is assumed to be continuous on (0, 1) by Axiom 4, we may conclude that $G(x)$ is constant over the reals of (0, 1). This implies that $g(x)$ is of the form $g(x)=cx^{\beta}$. This leads us to write in terms of $\tau(x)$ that

if $\beta=0$: $\tau(x)=\delta_1 x\log x+\delta_2 x+\delta_3$,
is $\beta\neq 0, -1$: $\tau(x)=\delta_1\{x^{\beta+1}/[\beta(\beta+1)]\}+\delta_2 x+\delta_3$,
if $\beta=-1$: $\tau(x)=\delta_1\log x+\delta_2 x+\delta_3$,

where δ_i's are constants of integration. Therefore we write, for the case of $\beta\neq 0, -1$,

$$M(s(n);n)=\hat{f}\left[\delta_1\sum_{i=1}^{n}\frac{(s_i)^{\beta+1}}{\beta(\beta+1)}+\delta_2\sum_{i=1}^{n}s_i+n\delta_3;n\right]$$
$$=\hat{f}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{(ns_i)^{\beta+1}-1}{\beta(\beta+1)};n\right],$$

where $\hat{f}(\delta_1 z+\delta_2+n\delta_3;b)=\hat{f}\{(1/n)\,n^{\beta+1}z-1/[n\beta(\beta+1)];n\}$.

It is noted that the representation includes the cases of $\beta=0$ and $\beta=-1$ by l'Hospital's rule.[7] The case $\beta=0$ is Theil's entropy measure.

Finally note that since $(1/n)\sum_{i=1}^{n}[[ns_i]^{\beta+1}-1]/[\beta^2+\beta]$ clearly obeys the principle of population, as does the left hand side of (16) (by Axiom 7), $\hat{f}(\ ;n)$ must also obey the principle of population. Hence:

THEOREM 2. *Axioms 1–7 lead to the class of generalized entropy measures such that*

$$M(s;n)=f(\mu_{\beta}^{n}(s)),\tag{17}$$

where

$$\mu_{\beta}^{n}(s)=\frac{1}{n}\sum_{i=1}^{n}\frac{(ns_i)^{\beta+1}-1}{\beta(\beta+1)}\qquad(18)^{8}$$

and f is a monotone transformation.

---

[7] We obtain also the property that $(1/n)\sum_{i=1}^{n}((ns_i)^{\beta+1}-1)/\beta(\beta+1)\geqslant 0$ for any $s\in S^n$ and $\beta\in(-\infty,+\infty)$.

[8] After completion of our work, we found that Toyoda [12] proposed $[1-\sum(ns_i)^{\alpha}/n]/\alpha$ as an inequality measure.

One word on the Atkinson [1] class:

$$I^n_\epsilon(s) = 1 - \left[\frac{1}{n}\sum_{i=1}^{n}(ns_i)^{1-\epsilon}\right]^{1/(1-\epsilon)} \qquad \text{defined for } \varepsilon \geqslant 0. \qquad (19)^9$$

It is easily seen that $\mu^n_\beta(s)$ and $I^n_\epsilon(s)$ are ordinally equivalent to each other for the values of $\varepsilon = -\beta > 0$ (including $\varepsilon = 1 = -\beta$). It is also noted that for $\varepsilon = 0 = \beta$ the equivalence property disappears. $I^n_\epsilon(s)$ tends to the totally insensitive measure as $\varepsilon \to 0$, whereas $M_\beta(s; n)$ tends to Theil's entropy measure[10] as $\beta \to 0$. This is one of the important differences between the classes $I^n_\epsilon(s)$ and $\mu^n_\epsilon(s)$. Also for $\beta > 0$ there is no measure in $I^n_\epsilon(s)$ which corresponds to $I^n_\beta(s)$.

4. CHOICE OF β

Although the generalized entropy class is a good deal more restrictive than the class of inequality measures mentioned at the beginning of the last section, we have the problem that for every β there will exist a different ordering of the points in S^n by the measure $\mu^n_\beta(s)$. How may a "suitable" value of β be chosen? We shall attempt to answer this by introducing a concept of sensitivity of inequality measures. Recall that since T_{ik} represents the marginal effect of transferring income from i to k on $f(\mu^n_\beta(s))$, the MRSI T_{ik}/T_{jk} represents the transfer effect from i to k relative to that from j to k. If s_i, s_j, and s_k are all different, this gives our measure of sensitivity.

We examine this in the case of $M^n(s) = f(\mu^n_\beta(s))$ when the MRSI takes the form $T_{ik}/T_{jk} = \{(s_i)^\beta - (s_k)^\beta\}/\{(s_j)^\beta - (s_k)^\beta\}$.

Let $s_i > s_j > s_k$. Then T_{ik}/T_{jk} is always greater than unity. It implies that the inequality improving effect of transfers is greater the richer the source of the transfer is. Here we are comparing the effect of making at the cost of i with that of making the same person k richer at the cost of j. Note that if k is selected from a richer class, then k's income must approach j's income and thus the MRSI T_{ik}/T_{jk} must become large. This effect is measured in the elasticity form by

$$E(\beta) \equiv \frac{\partial h}{\partial s_k}\frac{s_k}{h} = \beta\left[\frac{1}{(s_j/s_k)^\beta - 1} - \frac{1}{(s_i/s_k)^\beta - 1}\right],$$

where $h \equiv T_{ik}/T_{jk}$. A sensitive measure has a high $E(\beta)$, and conversely for insensitive measures T_{ik}/T_{jk} would not change much whether the beneficiary

---

[9] For $\varepsilon = 1$, $I^n_\epsilon(s) = 1 - n\{\prod_{i=1}^n s_i\}^{1/n}$.

[10] $\mu^n_0(s) = \sum_{i=1}^n s_i \log ns_i$.

k is selected from a rich class or a poor class. A symmetric argument may be made when $s_i < s_j < s_k$. It seems reasonable (though perhaps not essential) that the measure should be extremely sensitive to whether the income of a very rich man or a very poor man is used to accommodate a particular change in the shares s_i and s_j. Hence we propose that the following may be used to select an appropriate value of β.

AXIOM 8 (Sensitivity). *The measure to be used should be most sensitive in terms of $E(\beta)$ when $s_i < s_j < s_k$, or $s_i > s_j > s_k$.*

We are then able to state:

THEOREM 3. *Axioms 1–8 lead to Theil's entropy measure $M(s; n) = f(\sum_{i=1}^{n} s_i \log ns_i)$, where f is a monotone transformation.*

The proof is relegated to Appendix B.

APPENDIX A: PROOF OF THEOREM 1

We shall suppose that the decomposition referred to in Axiom 5 has been made in such a way that the nth individual belongs to N_2. Let us then pick an arbitrary triple of individuals i, j, k ($\neq n$), and consider that all of them belong to N_1. This is always possible by use of Axiom 2 which allows us to permute income recipients.

Using (2) we may write, for $s \in$ relative interior of S^n, $m_i(z; n) - m_k(z; n) = T_{ki}(s)$ which can be rewritten as, via Axiom 5,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \{F(G(\tilde{s}^1), H(s^2), \sum s_j^1, \sum s_j^2) - F(G(s^1), H(s^2), \sum s_j^1, \sum s_j^2)\},$$

where $\tilde{s}^1$ is the subvector of N_1 corresponding to the ε-perturbation of s.

This perturbation does not affect $\sum s_j^1$ and s^2 in any way precisely because in $s_n = 1 - \sum_{j=1}^{n-1} z_j$ the two terms $s_i + \varepsilon$ and $s_k - \varepsilon$ cancel out, though s_n itself is an element of s^2. The above formula is therefore reduced to $m_i(z; n) - m_k(z; n) = (\partial F/\partial G)[\partial G/\partial z_i - \partial G/\partial z_k]$. Likewise we have $m_j(z; n) - m_k(s; n) = (\partial F/\partial G)[\partial G/\partial z_j - \partial G/\partial z_k]$. Summarizing we have

$$[m_i(z; n) - m_k(z; n)]/[m_j[z; n) - m_k(z; n)]$$
$$= \left[\frac{\partial G}{\partial z_i} - \frac{\partial G}{\partial z_k}\right] \Big/ \left[\frac{\partial G}{\partial z_j} - \frac{\partial G}{\partial z_k}\right] \qquad (i, j, k) \subset N_1. \qquad \text{(A.1)}$$

The denominator in the above does not vanish in the light of Axiom 3 when $s_j \neq s_k$. Since the set of s with all its elements differing one another is dense in S, the results obtained under the imposition of $s_j \neq s_k$ can be extended to the whole of the relative interior of S.

Since G has as arguments only the variables pertaining to N_1, we have

$$\frac{m_{i\nu} - m_{k\nu}}{m_i - m_k} = \frac{m_{j\nu} - m_{k\nu}}{m_j - m_k} = \phi^\nu, \qquad \text{say, where } m_{i\nu} = \partial m_i / \partial z_\nu. \tag{A.2}$$

Since i, j, k and ν can be permuted by Axiom 2, (A.2) can be interpreted as holding for any quadruple extracted from $N - \{n\}$.

We shall suppose next a decomposition such that $\{n, i, j\} \subset N_1$ and $\{k, \nu\} \subset N_2$. We then have

$$\frac{m_{ik} - m_{i\nu}}{m_i} = \frac{m_{jk} - m_{j\nu}}{m_j}. \tag{A.3}$$

By (A.2) and (A.3) with the commutativity we may now put

$$\delta(z) \equiv \frac{\phi^k}{m_k}; \qquad \phi^k \equiv (m_{ik} - m_{jk})/(m_i - m_j), \tag{A.4}$$

where $k = 1, \ldots, n-1$ with arbitrary i and j $(\neq k)$.

Since $\partial \phi^k / \partial z_\nu = \partial \phi^\nu / \partial z_k$, (A.4) leads to

$$\delta_\nu / m_\nu = \delta_k / m_k. \tag{A.5}$$

The partial derivatives of $\delta(z)$ and $m(z)$ are proportional, and therefore by Lemma 1 of Goldman and Uzawa [5, p. 388], there exists a scalar-to-scalar transformation, say $W(\ )$, such that $\delta(z) = W(m(z; n))$.

Let x be a scalar and define[11] a function $Q(x)$ by

$$Q(x) = \int^x \exp\left[-\int^v W(\xi)\, d\xi\right] dv \qquad \text{or} \qquad Q''(x) + Q'(x)\, W(x) = 0.$$

Let $q(z) = Q(m(z; n))$. Then we have

$$\begin{aligned} q_{ik} - q_{ij} &= Q'' m_i (m_k - m_j) + Q'(m_{ij} - m_{ik}) \\ &= m_i (m_k - m_j)[Q'' + Q' W(m(z; n))] = 0. \end{aligned}$$

Thus $q_{jk} = q_{ij} = q_{ji} = q_{ji} \equiv \beta(z)$. Then $\beta_k = q_{ijk} = q_{ikj} = \beta_j$ by the commutativity (Axiom 4). Again by Lemma 1 of Goldman and Uzawa [5], we can establish that there exists a scalar-to-scalar function, say $\alpha(\ )$, such that

---

[11] See Samuelson [9, p. 178].

$$\beta(z) = \alpha(x); \qquad x = \sum_{i=1}^{n-1} z_i.$$

Let us then successively define $a(x) \equiv \int^x \alpha(v)\,dv$, $b(x) \equiv \int^x a(v)\,dv$, $c(z) \equiv b(\sum_{i=1}^{n-1} z_i)$, $p(z) \equiv q(z) - c(z) \equiv Q(m(z;n)) - c(z)$, for which we shall have $p_{ij}(z) = \beta(z) - \beta(z) = 0$ for all i and j $(i \neq j)$. It therefore follows that $Q(m(z;n)) - c(z)$ has to take an additive form, say,

$$Q(m(z;n)) - c(z) = h^1(z_1) + \cdots + h^{n-1}(z_{n-1}).$$

This leads us to write

$$\begin{aligned} Q(m(z;n)) &= Q(M(s;n)) \\ &= h^1(s_1) + \cdots + h^{n-1}(s_{n-1}) + h^n(s_n), \end{aligned}$$

where $h^n(s_n) = c(z) = b(\sum_{i=1}^{n-1} z_i) = b(1 - s_n)$.

Taking the inverse of Q, we may write

$$M(s;n) = f[h^1(s_1) + \cdots + h^{n-1}(s_{n-1}) + h^n(s_n)].$$

Finally the anonymity axiom (Axiom 2) implies that all the $h^i(\ )$'s are the same function. It goes without saying that f and h^i's here may depend upon the parameter n as in the statement of Theorem 1.

APPENDIX B: PROOF OF THEOREM 3

It suffices to show that max $E(\beta)$, $\beta \in (-\infty, +\infty)$, is attained at $\beta = 0$. We assume $s_k < s_j < s_i$. The other case may be proved analogously. Let $E(\beta) = \beta[1/(\eta^\beta - 1) - 1/(\xi^\beta - 1)]$, where $1 < \eta = (s_j/s_k) < s_i/s_k = \xi$. Then $E'(\beta) = H(\xi) - H(\eta)$, where $H(x) = \beta x^\beta \log x/(x^\beta - 1)^2 - 1/(x^\beta - 1)$. We shall have $H'(x) < 0$ for $\beta > 0$ and $H'(x) > 0$ for $\beta < 0$, and thus $E'(\beta) < 0$ for $\beta > 0$ and $E'(\beta) > 0$ for $\beta < 0$.

REFERENCES

1. A. B. ATKINSON, On the measurement of inequality, *J. Econ. Theory* **2** (1970), 244–263.
2. F. A. COWELL AND K. KUGA, Inequality measurement: An axiomatic approach, *European Econ. Rev.* **15** (1981), 287–305.
3. H. DALTON, The measurement of the inequality of incomes, *Econ. J.* **30** (1920), 348–361.
4. O. ÉLTETŐ AND E. FRIGYES, New income inequality measures as efficient tools for causal analysis and planning, *Econometrica* **36** (1968), 383–396.

5. S. M. Goldman and H. Uzawa, A note on separability in demand analysis, *Econometrica* **32** (1964), 387–398.
6. S.-C. Kolm, Unequal inequalities I, *J. Econ. Theory* **12** (1976), 416–442.
7. S.-C. Kolm, Unequal inequalities II, *J. Econ. Theory* **13** (1976), 82–111.
8. A. C. Pigou, "Wealth and Welfare," Macmillan & Co., London, 1912.
9. P. A. Samuelson, "Foundations of Economic Analysis," Harvard Univ. Press, Cambridge, Mass., 1947.
10. N. Takayama, Factors governing the inequality of income and money assets distributions in Japan—Shyotoku kinyushisan bunpu no fubyodō to sono yōin, *Keizai Kenkyu* 27 (1976), 134–142 (in Japanese).
11. H. Theil, "Economics and Information Theory," Chap. 4, North-Holland, Amsterdam, 1967.
12. T. Toyoda, Inequality of income distribution, *Kokumin Keizai* (1975), 15–41 (in Japanese).

JOURNAL OF ECONOMIC THEORY 25, 144–151 (1981)

Equilibria, the Core and Jurisdiction Structures in Economies with a Local Public Good: A Correction*

MYRNA WOODERS

Department of Economics, State University of New York,
Stony Brook, New York 11794

Received October 29, 1979

Shortly after the publication of the August 1978 issue of the *Journal of Economic Theory*, containing [4], I had occasion to reread the proof of one of the subsidiary theorems of the paper (Theorem 2). To my acute chagrin, I discovered some flaws in the proof of that theorem (these do not affect the other four theorems). The purpose of this note is to bring these to the attention of readers of the paper and to provide an existence theorem for the equilibrium in the several-private-goods case.

Since [4] was written, however, this author has developed an approximate Tiebout-type equilibrium concept for local public good economies. In [5], it is shown, for the one-private-good, one-local-public-good case, that the approximate equilibrium exists for all sufficiently large economies. The approximate equilibrium is more interesting than the exact equilibrium since number-theoretic assumptions are not required for existence. The nature of the difficulties with the existence "proof" in [4] and the results of [5] suggest that existence of the approximate equilibrium, in the more-than-one-private-good case, might not require the additional assumptions used to prove existence in this paper.

THEOREM 2′. *Assume the conditions stated in Theorem* 2 *of* [4]. *In addition, assume that*

(a) $u^t(n, x, y) = \psi(n)\, u^t(1, x, y)$ *for some function* $\psi(n)$ *and for each* t;

(b) $Y_0[n] = \{(\hat{x}, n\hat{z}): (\hat{x}, \hat{z}) \in Y_0[1]\}$.

Then the local public good economy has an equilibrium.

Condition (a) implies that the indifference map of the consumer in $(x, y$

---

* The errors of the article which necessitated this note and any errors in this note are entirely due to this author.

0022-0531/81/040144–08$02.00/0
Copyright © 1981 by Academic Press, Inc.

space does not change when n changes, although the utility levels realized from a given bundle of goods *might* change.

Condition (b) is that the per-capita imputs required to produce a given level of the local public good are unaffected by n. This assumption has some tradition in the theory of local public goods (see, for example, [3]).

These conditions together imply that the distinguished numbers for each type are independent of prices for private goods, i.e., $D(T; p) = D(T; p')$ for each T and for all $p, p' \in P$.

Proof. The strategy of the proof is, first, to determine the equilibrium prices for private goods, p^*, and then to construct prices for the local public good from p^* and problem (*).

The proof of the existence of equilibrium prices for private goods is essentially the same as that of the existence of equilibrium in Debreu [1]. However, the proof differs substantially from that of Debreu in that we determine a jurisdiction structure and then consider private-goods demand correspondences of jurisdictions (rather than of individual consumers). Since we aggregate out the local public good side of the jurisdictional demand correspondences, the correspondences considered and the entities constructed have different economic interpretations than the mathematically similar correspondences in [1].

Since free disposal was not assumed in [4], so that Debreu's proof can be applied, we first modify Y_1, the private goods production set. Let Y_1^c denote the convex hull of $Y_1 U - \Omega^K$. We say that a state of the economy, $s(\mathscr{J}) = \langle (n^N, x^N, y^N; \mathscr{J}), ((\hat{x}, \hat{z}; \mathscr{J}), z) \rangle$, is Y_1^c*-feasible* if $(n^N, x^N, y^N; \mathscr{J})$ is an allocation for $\mathscr{J}$ of N, $(\hat{x}, \hat{z}; \mathscr{J}) \in Y_0[\mathscr{J}]$ and $z \in Y_1^c$, where $\sum_{r=1}^{R} \hat{z}_r + z = \sum_{i \in N} (y^i - w^i)$. Informally, a state of the economy is feasible for Y_1^c if it is a feasible state of the economy for an economy where $Y_1 = Y_1^c$. It is easily verified that Y_1^c is a closed, convex cone, and it will be shown that $Y_1^c \cap \Omega^K = \{0\}$.

The proof will now proceed through several steps. In Step 1, an appropriate compact cube C is constructed. In Step 2, a private-goods production correspondence and a profit function are defined. In Step 3, a modification of problem (*) is investigated. In this modification, profits from private-goods production are added to the budget constraint ((1) of problem (*)). These profits are divided equally among consumers and later are shown to be zero in equilibrium. Also, private good usage is restricted to C. The aggregate jurisdisctional demand correspondence (for private goods) is defined in Step 4. It is shown to be convex-valued and u.h.c. (upperhemicontinuous). In the next step, the total excess demand correspondence for private goods, $E(p)$, is defined and it is shown that there is a p^* such that $p^* \cdot E(p^*) \leqslant 0$. The remaining steps demonstrate that p^* is an equilibrium price system for private goods and conclude the proof.

1. CONSTRUCTION OF C

First, it is necessary to show that $Y_1^c \cap \Omega^K = \{0\}$. Suppose there is a $z \in Y_1^c \cap \Omega^K$, where $z \neq 0$. From Carathéodory's theorem (see [2, p. 37]), there exists $z_1, \ldots, z_{K+1} \in Y_1 \cup -\Omega^K$ and $\lambda_1, \ldots, \lambda_{K+1} \in \Omega^1$, where $\sum_{k=1}^{K+1} \lambda_k = 1$ and $\sum_{k=1}^{K+1} \lambda_k z_k = z$. For each $k \in \{1, \ldots, K+1\}$, let $z'_k = z_k$ if $z_k \in Y_1$ and $z'_k = 0$ if $z_k \notin Y_1$. Then $\sum_{k=1}^{K+1} \lambda_k z'_k \in Y_1^c$ and $\sum_{k=1}^{K+1} \lambda_k z'_k \geqslant z$. Since $0 \in Y_1$, $z'_k \in Y_1$ for all k. Since Y_1 is convex, $\sum_{k=1}^{K+1} \lambda_k z'_k \in Y_1$ and, since $\sum_{k=1}^{K+1} \lambda_k z'_k \geqslant z$, we have a contradiction to the assumption that $Y_1 \cap \Omega^K = \{0\}$. Therefore $Y_1^c \cap \Omega^K = \{0\}$.

Define X by $X = \{v \in E^K$: for some jurisdiction structure $\mathscr{J} = \{J_1, \ldots, J_R\}$, there is a Y_1^c-feasible state of the economy, say $s(\mathscr{J}) = \langle (n^N, x^N, y^N; \mathscr{J}), ((\hat{x}, \hat{z}; \mathscr{J}), z) \rangle$, where $\sum_{i \in N} y^i - \sum_{r=1}^{R} \hat{z}_r = v\}$. Observe that $X = \{v \in \Omega^K$: $v = z + \sum_{i \in N} w^i$ for some $z \in Y_1^c\}$ so $X = (Y_1^c + \{\sum_{i \in N} w^i\}) \cap \Omega^K$. Since the asymptotic cone of $Y_1^c + \{\sum_{i \in N} w^i\}$ is Y_1^c and $Y_1^c \cap \Omega^K = \{0\}$, X is bounded (see [1, p. 23]).

Let C be a closed cube of E^K with center zero containing X in its interior.

Before concluding this section, we show that X and $Y_1^c \cap C$ are closed. To show that Y_1^c is closed, since C is closed, we need only show that Y_1^c is closed. From Carathéodory's theorem and the observation that Y_1 (and $-\Omega^K$) are both convex cones containing $0 \in \Omega^K$, it follows that $Y_1^c = Y_1 - \Omega^K$. Since Y_1 and, $-\Omega^K$ are positively semi-independent and closed, $Y_1 - \Omega^K$ is closed (see [1, pp. 22–23]) and $Y_1^c = Y_1 - \Omega^K$ is closed. Since Y_1^c is closed, X is the intersection of two closed sets and is therefore closed.

2. THE DEMAND CORRESPONDENCE AND PROFIT FUNCTION FOR PRIVATE-GOODS PRODUCTION

Define P^* by $P^* = \{p \in \Omega^K: \sum_{k=1}^{K} p_k = 1\}$; P^* is the closure of P.

Given $p \in P^*$, define $z(p)$ by $z(p) = \{z \in Y_1^c \cap C: p \cdot z \geqslant p \cdot z'$ for all $z' \in Y_1^c \cap C\}$; $z(p)$ is the truncated demand correspondence for private-goods production. Since C is compact and Y_1^c is closed, $Y_1^c \cap C$ is compact and $z(p) \neq \varnothing$ for any $p \in P^*$. As in Debreu [1], $z(p)$ is an u.h.c. correspondence from P^* to E^K. Also, $z(p)$ is convex-valued.

Given $p \in P^*$, define $\pi(p)$ by $\pi(p) = p \cdot z(p)$; $\pi(p)$ is the profit function. As in Debreu, $\pi(p)$ is a continuous function.

3. A MODIFICATION OF PROBLEM (*)

Consider Problem 1 below:

Maximize $u_n^t(x, y)$ subject to $(x, y, \hat{z}) \in \kappa(n; p)$ with respect to x, y, and $\hat{z}$, where $\kappa(n; p)$ is defined by $\kappa(n; p) = \{(x, y, \hat{z})$: (1), (2), (3), and (4) below are satisfied by $(x, y, \hat{z})\}$.

(1) $p \cdot (ny - \hat{z}) \leqslant p \cdot nw^t + n\pi(p)/m$,
(2) $(x, \hat{z}) \in Y_0[n]$,
(3) $(x, y) \geqslant 0$,
(4) $ny - \hat{z} \in C$.

Observe that for each $p \in P^*$, $\kappa(n; p)$ is compact, convex, and non-empty. Also, it can be shown by standard methods that $\kappa(n; p)$ is a continuous correspondence defined on P^*.

Since $u_n^t(x, y)$ is a continuous function defined on the compact set $\kappa(n; p)$, Problem 1 has a solution for any $p \in P^*$.

Let $\delta^T(n; p) = \{(x', y', \hat{z}') \in \kappa(n; p)$: $u_n^t(x', y') \geqslant u_n^t(x, y)$ for all $(x, y, \hat{z}) \in \kappa(n; p)\}$. Define $\bar{u}^t(n; p)$ by $\bar{u}^t(n; p) = u_n^t(x, y)$, where, for some $\hat{z}$, $(x, y, \hat{z}) \in \delta^T(n; p)$. It follows from the application of the Maximum Theorem ([1, p. 19]) that $\bar{u}^t(\cdot, \cdot)$ is a continuous function of p and $\delta^T(\cdot, \cdot)$ is a compact-valued and u.h.c. correspondence.

We now show that $\delta^T(n; p)$ is convex. Let $(x_1, y_1, \hat{z}_1)$, $(x_2, y_2, \hat{z}_2) \in \delta^T(n; p)$, let $\lambda \in (0, 1)$ and let $(x', y', \hat{z}') = \lambda(x_1, y_1, \hat{z}_1) + (1 - \lambda)(x_2, y_2, \hat{z}_2)$. Since $Y_0[n]$ is convex, $(x', \hat{z}') \in Y_0[n]$. Also, $p \cdot (ny' - \hat{z}') \leqslant p \cdot nw^t + n\pi(p)/m$. Since $ny_1 - \hat{z}_1 \in C$, $ny_2 - \hat{z}_2 \in C$ and C is convex, $ny' - \hat{z}' \in C$ so all the constraints of Problem 1 are satisfied by x', y' and $\hat{z}'$. Since $u^t(n \cdot, \cdot)$ is quasi-concave, $u^t(n, x', y') \geqslant u^t(n, x_1, y_1)$ but $u^t(n, x_1, y_1)$ is a maximum of $u^t(n \cdot, \cdot)$ subject to the constraints of Problem 1, so $u^t(n, x', y') = u^t(n, x_1, y_1)$. Therefore, $(x', y', \hat{z}') \in \delta^T(n, p)$.

4. THE AGGREGATE JURISDICTIONAL DEMAND CORRESPONDENCE FOR PRIVATE GOODS

First, observe that under the assumptions of this theorem, $u_n^t(x, y) = \psi(n)\, u^t(1, x, y)$ so n^* maximizes $\bar{u}^t(n; p)$ if and only if n^* maximizes $\psi(n)$. Consequently, let $D(T) = \{n \in I$: $\bar{u}^t(n; p) = u^{*t}(p)\}$; $D(T)$ represents the set of distinguished numbers of type T and is independent of $p \in P$.

From the assumption that for each $T \in \{A, B\}$, $|T| \geqslant \max_{p \in P} \{(n' - 1)(n'' - 1)$: n', $n'' \in D(T; p)\}$, we can select a juristiction structure, say $\mathscr{J} = \{J_1, \ldots, J_R\}$, so that $J_r \subseteq A$ or $J_r \subseteq B$ and, when $J_r \subseteq T$, $|J_r| \in D(T)$. This jurisdiction structure remains fixed throughout the remainder of this proof.

Define $\eta^T(p)$ by $\eta^T(p) = \{\sum_{J_r \subseteq T} (|J_r|\, y_r - \hat{z}_r)$: for some x, $(x, y_r, \hat{z}_r) \in \delta^T(|J_r|; p)\}$; $\eta^T(p)$ is the aggregate jurisdictional demand for private goods for jurisdictions whose membership is of type T.

Since $\delta^T(|J_r|; p)$ is convex for all r, $\eta^T(p)$ is convex.

Since $\delta^T(|J_r|, \cdot)$ is u.h.c., $\eta^T(p)$ is the composition of the Cartesian product of u.h.c. correspondences and a continuous function so $\eta'(\cdot)$ is u.h.c.

Let $\eta(p) = \eta^A(p) + \eta^B(p)$; $\eta(p)$ is the aggregate jurisdictional demand correspondence for private goods. It is immediate that $\eta(p)$ is a convex-valued, u.h.c. correspondence defined on P^*.

5. DEFINITION AND PROPERTIES OF THE AGGREGATE EXCESS DEMAND CORRESPONDENCE, $E(p)$

Define $E(p)$ by $E(p) = \eta^A(p) + \eta^B(p) - z(p) - \sum_{i \in N} w^i$; $E(p)$ is the aggregate excess demand correspondence for private goods. Since $\eta^A(p)$, $\eta^B(p)$ and $z(p)$ are non-empty, and convex-valued for all $p \in P^*$ and are all u.h.c. correspondences defined on P^*, $E(p)$ is non-empty for all $p \in P^*$ and a convex-valued, u.h.c. correspondence. Also, note that $E(p) \subset mC - Y_1^c \cap C - \{\sum_{i \in N} w^i\}$.

The set $mC - Y_1^c \cap C - \{\sum_{i \in N} w^i\}$ is compact as a sum of compact sets. Therefore, the conditions of the Debreu, Gale, Kuhn and Nikaido Lemma ([1, p. 82]) are satisfied and there exists $p^* \in P^*$ such that $E(p^*) \cap -\Omega^K \neq \emptyset$.

6. E-EQUILIBRIUM CONSUMPTIONS AND PRODUCTION

Let $v^* \in E(p^*) \cap -\Omega^K$. Since $v^* \in E(p^*)$, there is a $(x_r^*, y_r^*, \hat{z}_r^*) \in \delta^T(|J_r|; p^*)$ for each $J_r \subseteq T$ and for each T so that $\sum_{r=1}^R (|J_r|\, y_r^* - \hat{z}_r^*) \in \eta^A(p^*) + \eta^B(p^*)$. Also there is a $z' \in z(p)$ so that $\sum_{r=1}^R (|J_r|\, y_r^* - \hat{z}_r^*) - z' - \sum_{i \in N} w^i = v^*$. The set Y_1^c is convex, closed, and $-\Omega^K \subseteq Y_1^c$ so $z' \in Y_1^c$ and $v^* \in -\Omega^K$ implies that $z' + v^* \in Y_1^c$.

Let $z^* = z' + v^*$ so $\sum_{r=1}^R (|J_r|\, y_r^* - \hat{z}_r^*) - z^* - \sum_{i \in N} w_i = 0$. Let $s^*(\mathcal{A}) = \langle (n^{*N}, x^{*N}, y^{*N}; \mathcal{A}), ((x^*, z^*; \mathcal{A}), z^*) \rangle$ denote the state of the economy where, when $i \in J_r \subseteq T$, $y^{*i} = y_r^*$, $x^{*i} = x_r^*$ and $n^{*i} = |J_r|$ and where $(\hat{x}^*, \hat{z}^*; \mathcal{A}) = ((\hat{x}_1^*, \hat{z}_1^*), \ldots, (\hat{x}_R^*, \hat{z}_R^*))$. Since $\sum_{r=1}^R (|J_r|\, y_r^* - \hat{z}_r^*) - z^* - \sum_{i \in N} w^i = 0$, $s^*(\mathcal{A})$ is a Y_1^c-feasible state of the economy. Consequently, $(|J_r|\, y_r^* - \hat{z}_r^*)$ and z^* are in the interior of the cube C for all r.

7. PROPERTIES OF z^*

Since $(|J_r|\, y_r^* - \hat{z}_r^*)$ is in the interior of the cube C and utility functions are monotonic, it follows that $p^* \gg 0$ and $p^* \cdot (|J_r|\, y_r^* - \hat{z}_r^*) = p^* \cdot |J_r|\, w' + |J_r|\, \pi(p^*)/m$ when $J_r \subseteq T$. Consequently, $p^* \cdot \sum_{r=1}^R (|J_r|\, y_r^* - \hat{z}_r^*) - p^* \cdot z' - p^* \cdot \sum_{i \in N} w^i = 0$ since $\pi(p^*) = p^* \cdot z'$, so $p^* \cdot v^* = 0$.

Since $p^* \gg 0$, $v^* \leqslant 0$ and $p^* \cdot v^* = 0$, $v^* = 0$ and $z^* = z'$. Since $z^* = z'$, $z^* \in z(p^*)$ and $p^* \cdot z^* \geqslant p^* \cdot z$ for all $z \in Y_1^c \cap C$. Since z^* is in the interior of C, $p^* \cdot z^* \geqslant p^* \cdot z$ for all $z \in Y_1^c$. Since $Y_1^c = Y_1 - \Omega^K$, there is a $z_1 \in Y_1$ and $z_2 \in -\Omega^K$ so that $z_1 + z_2 = z^*$. Since $p^* \gg 0$ if $z_2 \neq 0$, $p^* \cdot z_1 > p^* \cdot z^*$; which is a contradiction, so $z_2 = 0$ and $z^* \in Y_1$. Also, $p^* \cdot z^* \geqslant p^* \cdot z$ for all $z \in Y_1$ since $Y_1 \subseteq Y_1^c$. Since Y_1 is a closed, convex cone, $p^* \cdot z^* = \pi(p^*) = 0$.

8. PROPERTIES OF E-EQUILIBRIUM CONSUMPTIONS

Since $p^* \gg 0$, $p^* \in P$. Since $\pi(p^*) = 0$, when $t \in J_r \subseteq T$, $p^* \cdot (|J_r|\, y_r^* - \hat{z}_r^*) = p^* \cdot |J_r|\, w^t$ and x_r^*, y_r^*, and $\hat{z}_r^*$ are solution values for problem $(*)$ for x, y, and $\hat{z}$, respectively. Since $|J_r| \in D(T) = D(T; p^*)$, $u^t(|J_r|, x_r^*, y_r^*) = u^{*t}(p^*)$.

9. CONSTRUCTION OF $\mathscr{J}$-EXTENDED PRICE AND PROFIT SYSTEMS FOR LOCAL PUBLIC GOODS

For each $n \leqslant m$, select $(\hat{x}, \hat{z}) \in$ Boundary of $Y_0[n]$, where $\hat{x} > 0$. Let $Y'[n] = \{(\hat{x}', p^* \cdot \hat{z}') \colon (\hat{x}', \hat{z}') \in Y_0[n]\}$. It is easy to verify that $Y'[n]$ is a closed convex cone. From the Minkowski Separating Hyperplane theorem, there are numbers $\gamma_1(n)$, $\gamma_2(n)$ (not both zero) and c so that $\gamma_1(n)\hat{x} + \gamma_2(n)p^* \cdot \hat{z} = c \geqslant \gamma_1(n)\hat{x}' + \gamma_2(n)p^* \cdot \hat{z}$ for all $(\hat{x}', \hat{z}') \in Y_0[n]$. Since $0 \in Y'[n]$, $c \geqslant 0$. Since $Y'[n]$ is a cone containing $(\hat{x}, \hat{z})$ where $\hat{x} > 0$ and $\hat{z} < 0$, $c = 0$, and both $\gamma_1(n)$ and $\gamma_2(n)$ are non-zero. Since $\hat{x} \in \Omega^1$ and $\hat{z} \in -\Omega^K$, $\gamma_1(n)$ and $\gamma_2(n)$ have the same sign. Define $\gamma(n)$ by $\gamma(n) = \gamma_1(n)/\gamma_2(n)$; $\gamma(n) > 0$. Then $\gamma(n)\hat{x} + p^* \cdot \hat{z} = 0 \geqslant \gamma(n)\hat{x}' + p^* \cdot \hat{z}'$ for all $(\hat{x}', \hat{z}') \in Y_0[n]$.

Define $\mathscr{U}[\mathscr{J}]$ as in [4, p. 337]; $U[\mathscr{J}] = (|J_1|, \ldots, |J_r|, V_{R+1}, \ldots, V_{R+t}) = (V_1, \ldots, V_u, \ldots, V_{R+t})$; $U[\mathscr{J}]$ is a vector of jurisdiction and quasi-jurisdiction sizes. Define $\gamma(\mathscr{J})$ by $\gamma(\mathscr{J}) = (\gamma(|J_1|), \ldots, \gamma(|J_R|), \gamma(V_{R+1}), \ldots, \gamma(V_{R+t})) = (\gamma_1, \ldots, \gamma_u, \ldots, \gamma_{R+t})$. Define $\pi(\mathscr{J})$ by $\pi(\mathscr{J}) = (\pi_1, \ldots, \pi_u, \ldots, \pi_{R+t}) = 0 \in \Omega^{R+t}$.

10. A COMPETITIVE LOCAL PUBLIC EQUILIBRIUM

We conclude the proof by showing that $e = \langle s^*(\mathscr{J}), p^*, \gamma(\mathscr{J}), \pi(\mathscr{J})\rangle$, where $s^*(\mathscr{J}) = \langle (n^{*N}, x^{*N}, y^{*N}; \mathscr{J}), (\hat{x}^*, \hat{z}^*; \mathscr{J}), z^*\rangle$, is a competitive local public equilibrium.

We have already shown that e satisfies (i) and (ii) of the definition of the equilibrium.

Also, in Step 9 it was shown that $0 = \pi_u = \max\{\gamma_u \hat{x}'_u + p^* \cdot \hat{z}'_u : (\hat{x}'_u, \hat{z}'_u) \in Y_0(V_u)\}$. We now will show that for all $r \in \{1, \dots, R\}$, $\pi_r = \gamma_r \hat{x}^*_r + p^* \cdot \hat{z}^*_r$. Suppose $\pi_r > \gamma_r \hat{x}^*_r + p^* \cdot \hat{z}^*_r$. Recall that $p^* \cdot (|J_r|\, y^*_r - \hat{z}^*_r) = |J_r|\, p^* \cdot w^i$ when $J_r \subseteq T$ and $(\hat{x}^*_r, \hat{z}^*_r) \in Y_0\|J_r\|$. Let $(\hat{x}', \hat{z}') \in Y_0\|J_r\|$, where $\hat{x}' > 0$ and $\gamma_r \hat{x}' + p^* \cdot \hat{z}' = 0$. Since $Y_0\|J_r\|$ is a closed, convex cone, there is a real number $\lambda \in \Omega$ so that $\hat{x}^*_r = \lambda \hat{x}'$. Then $(\hat{x}^*_r, \lambda \hat{z}') \in Y_0\|J_r\|$ and $\gamma_r \hat{x}^*_r + p^* \cdot \lambda \hat{z}' = 0$. Since $0 > \gamma_r \hat{x}^*_r + p^* \cdot \hat{z}^*_r$, $p^* \cdot \hat{z}^*_r < p^* \cdot \lambda \hat{z}'$. Consequently, there is a $y' \geqslant y^*_r$ so that $p^* \cdot (|J_r|\, y' - \lambda z') = |J_r|\, p^* \cdot w^i$. Since $\bar{u}^i(|J_r|, p^*) = u^i(|J_r|, x^*_r, y^*_r) < u^i(|J_r|, x^*_r, y')$ from monotonicity, $p^* \cdot (|J_r|\, y' - \lambda \hat{z}') = |J_r|\, p^* \cdot w^i$, and $(x^*_r, \lambda \hat{z}') \in Y_0\|J_r\|$, we have a contradiction to the conclusion that x^*_r, y^*_r, z^*_r are solution values for Problem (*). Therefore $\gamma_r \hat{x}^*_r + p^* \cdot \hat{z}^*_r = 0$ and (iii) of the definition of equilibrium is satisfied.

Since $\gamma_r \hat{x}^*_r + p^* \cdot \hat{z}^*_r = 0$ and, when $i \in J_r$, $p^* \cdot (y^{i*} - w^i) - p^* \cdot \hat{z}^*_r / |J_r| = 0$, $p^* \cdot (y^{i*} - w^i) + \gamma_r \cdot \hat{x}^*_r / |J_r| = 0$. Suppose $u^i(n', x', y') > u^i(n^{*i}, x^{*i}, y^{*i})$ where $n' = V_u$ for some $u \in \{1, \dots, R + v\}$ and $p^* \cdot (y' - w^i) + (\gamma_u / n') x' \leqslant 0$. Let $(\hat{x}, \hat{z}) \in Y_0|n'|$ so that $\hat{x} > 0$ and $\gamma_u \cdot \hat{x} + p^* \cdot \hat{z} = 0$. Since $\hat{x} > 0$ and $Y_0|n'|$ is a closed convex cone, there is a $\lambda \in \Omega^1$ so that $\lambda \hat{x} = x'$ and $(x', \lambda \hat{z}) \in Y_0|n'|$. Also $\gamma_u \cdot x' + p^* \cdot \lambda \hat{z} = 0$ so $p^* \cdot (y' - w^i) + p^* \cdot \lambda \hat{z} / n' \leqslant 0$. Then there is a $y'' \geqslant y'$ so that $p^* \cdot (y'' - w^i) + p^* \cdot \lambda \hat{z} / n' = 0$ and x', y'', and $\lambda \hat{z}$ satisfy the constraints of problem (*). Since $|J_r| \in D(T)$ when $i \in T$, $\bar{u}^i(|J_r|, p^*) = u^{*i}(p^*)$. But $u^i(n', x', y'') > \bar{u}^i(|J_r|, p^*)$ and x', y'' satisfy the constraints of problem (*) given p^* and n', which contradicts the conclusion that $u^i(|J_r|, x^*_r, y^*_r) = u^{*i}(p^*)$. Q.E.D.

The first major error in [4] in the statement of Lemma 2. The following version is correct.

LEMMA 2'. *Given $T \in \{A, B\}$ there exists a finite set of subsets of P, say $P^1, \dots, P^Q$, such that $\bigcup_{q=1}^{Q} P^q = P$ and for all q and all $p, p' \in P^q$, $D(T; p) = D(T; p')$. In addition, if $n \in D(T; p^l)$ for some sequence $p^l \in P^q$, and $p^l \to p \in P$, then $n \in D(T; p)$.*

From this lemma, it follows that for each P^q we can arbitrarily select one jurisdiction structure so that each agent of each type is contained in a jurisdiction of distinguished number size for type T relative to p for all $p \in P^q$.

The excess demand correspondence for each type ($\mathcal{A}^t(p)$ in [1]), given the selected jurisdiction structure for $p \in P^q$, is then convex valued and upper hemi-continuous on P^q. However, the excess demand correspondence created by selecting one jurisdiction structure for each P^q is not necessarily convex-valued on P (contrary to my claim in Lemma 4). The additional conditions on Theorem 2' were imposed to ensure u.h.c. and convex valuedness of $E(p)$

at the boundaries of P^q. These observations prompt the conjecture that the approximate equilibrium has much better existence properties.

REFERENCES

1. G. DEBREU, "Theory of Value," Wiley, New York, 1959.
2. W. HILDEBRAND, "Core and Equilibria of a Large Economy," Princeton Univ. Press, Princeton, N. J., 1974.
3. M. MCGUIRE, Private good clubs and public good clubs: Economic models of group formation, *Swedish J. Econ.* **74** (1972), 84–99.
4. M. WOODERS, Equilibria, the core and jurisdiction structures in economies with a local public good, *J. Econ. Theory* **18** (1978), 328–348.
5. M. WOODERS, The Tiebout hypothesis: Near optimality in large local public good economies, *Econometrica* **48** (1980), 1467–1485.

JOURNAL OF ECONOMIC THEORY 25, 152–162 (1981)

A Note on the Probability of Casting a Decisive Vote

GARY CHAMBERLAIN AND MICHAEL ROTHSCHILD*

Department of Economics, University of Wisconsin, Madison, Wisconsin 53706

Received April 8, 1980; revised July 22, 1980

Although it is widely believed that the probability that any one vote will decide an election is so small that the prospect of influencing an election cannot explain why rational people vote,[1] there have been to our knowledge few attempts to calculate the relevant probability. The calculations which we know of are neither general nor rigorous.[2]

The purpose of this note is to establish under fairly general circumstances that the probability that a voter casts the decisive ballot in an election in which $2N+1$ voter choose among two alternatives is of order N^{-1}. This is both smaller and larger than one might on first, or second thought, expect. An individual's vote matters if, and only if, the other $2N$ members of the polity split their votes evenly. The simplest way to calculate the probability of this happening is to suppose that the votes of individual members of society are independent binomial random variables while the probability that any individual votes yes is p. Then the probability that exactly N of the $2N$ members of society vote yes is $\binom{2N}{N} p^N (1-p)^N$, which we will write as $Q(N,p)$.

It is convenient to use Stirling's formula to evaluate $Q(N,p)$. Since $\binom{2N}{N} \approx 2^{2N}/\sqrt{\pi N}$, we see that for $p=1/2$,

$$Q(N, 1/2) \approx \frac{1}{\sqrt{\pi N}}.$$

---

* We thank John Ferejohn, Gerald Kramer and Steve Stigler for suggestions and the National Science Foundation for research support.

[1] See, for example, Ferejohn and Fiorina [4]. For some evidence that is consistent with the view that people are more likely to vote if the probability of their influencing the outcome is higher, see Barzel and Silverberg [1]. This evidence is, of course, not inconsistent with the position that the probability of influencing an election is so small that it cannot explain why rational men and women vote.

[2] We review this literature briefly in the Appendix.

0022-0531/81/040152–11$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

Since $(\sqrt{N})^{-1}$ decreases fairly slowly, the probability of having a decisive vote is, in this case, quite high. For example, in a society of 100,001 persons, the probability of having a decisive vote is approximately 0.0025; in a society of 1,000,001, this probability is about 0.0008.

However, if $p \neq 1/2$, $Q(N,p)$ decreases much more quickly.

$$Q(N,p) \approx \frac{2^{2N}p^N(1-p)^N}{\sqrt{\pi N}}$$
$$\approx Q(N,1/2)\exp N[2\log 2 + \log p + \log(1-p)].$$

If $p \neq 1/2$, $[2\log 2 + \log p + \log(1-p)] < 0$ and thus $Q(N,p)$ is of order e^{-cN} for some strictly positive c and goes to zero very quickly, as Table I shows.

What are we to make of these two results? One might argue that since the case $p = 1/2$ is unlikely ever to hold exactly, it must be that the probability that an individual's vote matters is in general of order e^{-cN}. It follows that it almost never makes sense to vote in any moderately large polity. We prefer to take seriously the notion that p is unknown and calculate

$$\int_0^1 \binom{2N}{N} p^N(1-p)^N\,dF(p) = \int_0^1 Q(N,p)\,dF(p). \tag{1}$$

One way of thinking about how (1) might arise is to suppose that nature picks a p at random from the distribution $F(p)$. Then $2N$ coins are tossed; each has probability p of coming up heads. For that p the probability of observing exactly N heads is $Q(N,p)$. The total probability of observing N

TABLE I

Probability of Having Decisive Vote as a Function of M (Size of Polity) and p (Probability a Polity Member Votes "Yes")

p \ M	1001	100.001	1.000.001	100.000.001
0.4	0.3×10^{-10}	*	*	*
0.45	0.2×10^{-3}	*	*	*
0.475	0.007	*	*	*
0.49	0.021	0.5×10^{-11}	*	*
0.4999	0.025	0.0025	0.0008	0.1×10^{4}
0.5	0.025	0.0025	0.0008	0.8×10^{-4}

* Less than 10^{-35}.

heads is gotten by integrating against $F(p)$ as in (1). A model in which the probability that each person votes yes is unknown but which does *not* give the expression (1) as the probability of a tie is the following: Each individual in the society has his own probability of voting yes. When the ith voter enters the polling booth, he votes yes with probability p_i. The p_i's are independently and identically distributed in the population according to the distribution function B. In this model, the probability that an individual votes yes is simply $\mu_B = \int_0^1 p\, dB(p)$ and since voters are independent, the probability that their N of $2N$ votes are yes is $Q(N, \mu_B)$. Thus the only important aspect of B is its mean, μ_B. If μ_B is known then the model is the same as one in which each voter's probability of voting yes is μ_B. If B is not known, then this argument shows that all that we need to know about B is its mean; the only important uncertainty about B is uncertainty about μ_B. This is the model discussed above which leads to (1).

In his original memoir, Bayes calculated $\int_0^1 Q(N,p)\, dF(p)$ for the special case where $F(p)$ is a uniform distribution. In this case $\int_0^1 Q(N,p)\, dF(p)$ reduces to $\int_0^1 \binom{2N}{N} p^N (1-p)^N\, dp$. This expression is easily evaluated by noting, as Bayes did, that if $X_0, X_1, \ldots, X_M$ is a random sample of $M+1$ random variables, each of which has a uniform distribution on $[0, 1]$, then if $X_0 = p$, $\binom{M}{k} p^k (1-p)^{M-k}$ is just the probability that k of the succeeding M random variables are less than p and $M-k$ of them are greater than p. Integrate over p to obtain the probability that in the sample, the first random variable observed is the $k+1$st smallest. Symmetry implies that this probability must be equal to $(M+1)^{-1}$, independent of k, or that

$$\int_0^1 \binom{M}{k} p^k (1-p)^{M-k}\, dp = (M+1)^{-1}.$$

It follows that $\int_0^1 \binom{2N}{N} p^N (1-p)^N\, dF(p)$ is equal to $(2N+1)^{-1}$ when $F(p)$ is uniform. We now show that this result is a prototype of the general case, that if one is uncertain about the probability that others vote yes then the probability of one's vote being decisive is of order N^{-1}.

PROPOSITION 1. *If the probability that voters will vote yes is distributed according to a distribution* $F(p)$ *with a continuous density* $f(p)$ *in the neighborhood of* 1/2, *then*

$$\lim_{N \to \infty} N \int_0^1 Q(N,p)\, dF(p) = (1/2) f(1/2).$$

Proof. Note that $N \int_0^1 Q(N,p)\, dF(p) = \int_0^1 N \binom{2N}{N} p^N (1-p)^N\, dF(p)$. Let $A = \{p \in (1/2 - \varepsilon, 1/2 + \varepsilon)\}$ and $B = \{p \in [0, 1/2 - \varepsilon] \cup [1/2 + \varepsilon, 1]\}$ for

$\varepsilon > 0$. Since $\binom{2N}{N} p^N(1-p)^N$ is uniformly of order e^{-cN}, $c > 0$, for $p \in B$, $N\binom{2N}{N} p^N(1-p)^N \to 0$ uniformly in p for $p \in B$. Thus,

$$\lim_{N\to\infty} \int_B N\binom{2N}{N} p^N(1-p)^N \, dF(p) = 0.$$

For any $\delta > 0$ we can choose ε small enough that $f(1/2) - \delta \leqslant f(p) \leqslant f(1/2) + \delta$ for $p \in A$. Hence

$$\begin{aligned}&\left[f\left(\frac{1}{2}\right) - \delta\right] \int_A N\binom{2N}{N} p^N(1-p)^N \, dp \\ &\leqslant \int_A N\binom{2N}{N} p^N(1-p)^N f(p)\, dp \\ &\leqslant \left[f\left(\frac{1}{2}\right) + \delta\right] \int_A N\binom{2N}{N} p^N(1-p)^N \, dp.\end{aligned}$$

From the uniform case, we know that

$$\lim_{N\to\infty} \int_0^1 N\binom{2N}{N} p^N(1-p)^N \, dp = \lim_{N\to\infty} \int_A N\binom{2N}{N} p^N(1-p)^N \, dp = \frac{1}{2}.$$

Hence

$$[f(\tfrac{1}{2}) - \delta]\, \tfrac{1}{2} \leqslant \lim_{N\to\infty} N\int_0^1 Q(N,p)\, dF(p) \leqslant [f(\tfrac{1}{2}) + \delta]\, \tfrac{1}{2}$$

for any $\delta > 0$, and so

$$\lim_{N\to\infty} N\int_0^1 Q(N,p)\, dF(p) = \tfrac{1}{2} f(\tfrac{1}{2}).$$

Note that Proposition 1—and its proof—remains true when $f(1/2) = 0$.

The result of Proposition 1 is considerably more general than the simple model from which it comes. Suppose that the polity consists of several distinct groups, and that we believe the propensity to vote yes varies from group to group. For simplicity we consider the case of two groups, but the analysis generalizes to any arbitrary number of groups. If voters come from two distinct groups, for concreteness think of men and women, then our uncertainty about the propensity to vote yes is expressed by a joint prior density $f(p_1, p_2)$ of the propensities for men and women to vote yes.

Information about propensities of different groups to vote is most easily thought of in terms of the marginal distributions for each group;

$f_1(p_1) = \int_0^1 f(p_1, p_2)\, dp_2$ and $f_2(p_2) = \int_0^1 f(p_1, p_2)\, dp_1$, and it is natural to ask why the propensities are not independent in which case

$$f(p_1, p_2) = f_1(p_1) f_2(p_2). \tag{2}$$

We consider the more general case where (2) does not necessarily hold. It is worthwhile to digress on the meaning of independence in this case. Consider an arbitrary sequence of voters. Suppose a statistician has observed the vote of N_1 men and N_2 women. He wants to use this information, and his prior information, $f(p_1, p_2)$, to predict the next man's vote. If he uses the information about the votes of women to make this prediction, then the prior distribution is not independent. We believe that, in general, votes of men will be used to predict the votes of women and vice-versa, and thus we consider the general case where (2) does not hold. This includes the case where voting behavior is believed to be completely independent of sex so that $f(p_1, p_2) = f(p_2, p_1)$.

We now consider the probability of a tie vote in an electorate of size $2N$, in which a fraction, α, of the voters are men and a fraction $(1 - \alpha)$ are women. Different models can be used to calculate this probability. The simplest supposes that the vote is a random sample of size $2N$, where each vote is chosen as follows. Before voting begins a pair (p_1, p_2) is drawn at random from the distribution $f(p_1, p_2)$. With probability α a man is chosen to vote first; he votes yes with probability p_1 and no with probability $(1 - p_1)$. If, as will happen with probability $1 - \alpha$, a woman is chosen to vote first, she votes yes with probability p_2 and no with probability $1 - p_2$. The process, which is one of sampling with replacement, continues in this way. At each stage the probabilities of men and women voting yes are given by the (p_1, p_2) selected initially. The kth vote in this sequence is a random variable X_k which assumes the values of 0 or 1. An electorate of size k produces a random sequence $X_1, \ldots, X_k$. It is a consequence of random sampling that the probability of any observed sequence of votes is independent of the order in which they are counted or that

$$\Pr\{X_1, \ldots, X_k\} = \Pr\{X_{\pi(1)}, \ldots, X_{\pi(k)}\}, \tag{3}$$

where π is any permutation of the first k integers. Since (3) holds for any k, the random variables $X_1, \ldots, X_m$ are what is known as *infinitely exchangeable*. It follows from a theorem of de Finetti's[3] that there is a cumulative distribution function G, independent of m, such that

$$\Pr\{X_1 = x_1, \ldots, X_m = x_m\} = \int_0^1 p^j (1 - p)^{m-j}\, dG(p), \tag{4}$$

---

[3] For a discussion of this theorem and the idea of exchangeability, see Diaconis [3].

where $j = \sum_{i=1}^{m} x_i = S_m$. It follows that

$$\Pr\{S_m = j\} = \int_0^1 \binom{m}{j} p^j (1-p)^{m-j}\, dG(p). \tag{5}$$

Since (5) holds for any m including $m = 1$, $G(p)$ must be the cumulative distribution function of the mixture $\alpha p_1 + (1-\alpha)p_2$. The density of $G(p)$ is

$$g(p) = \frac{1}{1-\alpha}\int_0^1 f\left(p_1, \frac{p-\alpha p_1}{1-\alpha}\right) dp_1. \tag{6}$$

Proposition 1 and the representations (4) and (5) immediately imply

PROPOSITION 2. *If voting is conceived of as drawing a sample of size* $2N$ *with replacement from a population in which the ratio of men to women is* $(\alpha/(1-\alpha))$ *and the prior probability of men and women voting yes is given by the continuous density* $f(p_1, p_2)$, *then* $N\Pr\{S_{2N} = N\} \to (1/2)g(1/2)$, *where* $g(p)$ *is given by* (6).

Thinking of voting as sampling with replacement is reasonable if the fraction of men and women who vote is not large. If turnout is high, the approximation is inaccurate. The sequence of votes generated by sampling without replacement will not in general be infinitely exchangeable so that de Finetti's theorem does not apply.[4]

Nonetheless, the conclusions of Propositions 1 and 2 continue to hold

---

[4] A sequence of length N which arose from sampling (without replacement) two populations of size $N_1 + N_2 = N$ would be exchangeable of order N but it would not in general be infinitely exchangeable. Suppose for example that the population consisted of one man and one woman and that men always voted yes while women always voted no. Assuming that the probability that the man votes first is 1/2, then

$$\Pr\{X_1 = 1, X_2 = 0\} = \Pr\{X_1 = 0, X_2 = 1\} = \tfrac{1}{2} \tag{7}$$

and

$$\Pr\{X_1 = 0, X_2 = 0\} = \Pr\{X_1 = 1, X_2 = 1\}. \tag{8}$$

If this sequence, which is exchangeable, could be extended to an infinitely exchangeable sequence, then there would exist a distribution function G such that (7) and (8) could be represented as in (4). But (8) implies

$$0 = \int_0^1 p^2\, dG(p) = \int_0^1 (1-p)^2\, dG(p),$$

or that the probability measure which G represents puts mass one at the point 0 and at the point 1, which is impossible. Diaconis's paper, which gives this example, also shows that it is not pathological.

when voting is explicitly modeled as sampling without replacement. Suppose that N_1 men and N_2 women vote, where $N_1 \leqslant N_2$ and $N_1 + N_2 = 2N$. Let R_1 and R_2 be the number of men and of women voting yes. Then, if the prior distribution for p_1 and p_2 has a density $f(p_1, p_2)$, the probability of a tie is

$$\Pr\{R_1 + R_2 = N\} = \int_0^1 \int_0^1 S_N(p_1, p_2) f(p_1, p_2)\, dp_1\, dp_2, \tag{9}$$

where

$$\begin{aligned} S_N(p_1, p_2) &= \Pr\{R_1 + R_2 = N \mid p_1, p_2\} \\ &= \sum_{n=0}^{N_1} \binom{N_1}{n} p_1^n (1-p_1)^{N_1 - n} \binom{N_2}{N-n} p_2^{N-n} (1-p_2)^{N_2 - N + n}. \end{aligned} \tag{10}$$

Proposition 3. *Suppose $N \to \infty$ in such a way that $N_1/2N \to \alpha$. Then if $f(p_1, p_2)$ is bounded and continuous, $N \Pr\{R_1 + R_2 = N\} \to 1/2g(1/2)$, where $g(p)$ is given by* (6).

Proof. Note that

$$\begin{aligned} \psi_{N,N_1}(p_1, p_2) &= (N_2 + 1)\, S_N(p_1, p_2) \\ &= \sum_{n=0}^{N_1} \frac{1}{N_1 + 1} f_\beta(p_1 \mid n+1, N_1 - n + 1) \\ &\quad \times f_\beta(p_2 \mid N - n + 1, N_2 - N + n + 1), \end{aligned} \tag{11}$$

where $f_\beta(p \mid A, B)$ is the density of a Beta distribution with parameters A and B. The important thing about (11) is that the right-hand side is a probability density for (p_1, p_2) with parameters N_1 and N. Since we are interested in taking limits where $N \to \infty$ and $N_1/2N \to \alpha$, we can without abuse ignore the dependence of ψ_{N,N_1} on N_1 and write $\psi_N(p_1, p_2)$. In this notation, the probability of a tie times N is

$$\begin{aligned} &N \int_0^1 \int_0^1 S_N(p_1, p_2) f(p_1, p_2)\, dp_1\, dp_2 \\ &\quad = N \int_0^1 \int_0^1 \frac{1}{N_2 + 1} \psi_N(p_1, p_2) f(p_1, p_2)\, dp_1\, dp_2 \\ &\quad = \frac{1}{2} \frac{2N}{N_2 + 1} \int_0^1 \int_0^1 f(p_1, p_2)\, \psi_N(p_1, p_2)\, dp_1\, dp_2 \\ &\quad = \frac{1}{2} \frac{2N}{N_2 + 1} E_{\psi_N} f(p_1, p_2), \end{aligned}$$

where the expectation is taken with respect to the distribution of (p_1, p_2) specified by ψ_N.

Observe that if p_1 and p_2 are distributed according to $\psi_N(p_1, p_2)$, the marginal distribution of p_1 is uniform on $[0, 1]$ since

$$\begin{aligned}
h(p_1) &= \int_0^1 \psi_N(p_1, p_2)\, dp_2 \\
&= \frac{1}{N_1+1} \sum_{n=0}^{N_1} f_\beta(p_1 | n+1, N_1-n+1) \\
&\quad \times \int_0^1 f_\beta(p_2 | N-n+1, N_2-N+n+1)\, dp_2 \\
&= \frac{1}{N_1+1} \sum_{n=0}^{N_1} f_\beta(p_1 | n+1, N_1-n+1) \\
&= \frac{1}{N_1+1} \sum_{n=0}^{N_1} \frac{(N_1+1)!}{n!\,(N_1-n)!} p_1^n (1-p_1)^{N_1-n} \\
&= \sum_{n=0}^{N_1} \binom{N_1}{n} p_1^n (1-p_1)^{N_1-n} = 1.
\end{aligned}$$

Now define the random variable $Z = \alpha p_1 + (1-\alpha) p_2$. Since p_1, p_2 are distributed according to the density ψ_N, the distribution of Z is determined by N, which we indicate by writing Z_N. Of the three random variables, p_1, p_2, and Z_N, the joint distribution of any two determines the third. Let H_N be the joint distribution function for (p_1, Z_N). Then

$$\int_0^1 \int_0^1 f(p_1, p_2)\, \psi_N(p_1, p_2)\, dp_1\, p_2 = \int_0^1 \int_0^1 f\left(p_1, \frac{z-\alpha p_1}{1-\alpha}\right) dH_N(p_1, z).$$

Suppose that if $N \to \infty$ (and, of course, $N_1/2N \to \alpha$), then the distribution of the random variable Z_N converges weakly to the distribution which puts probability one at $Z = 1/2$. Then $H_N(p_1, z)$ converges weakly to $H(p_1, z)$, where $H(p_1, z)$ puts probability one at $Z = 1/2$ and the marginal distribution of p_1 is uniform on $[0, 1]$. Then, since f is bounded and continuous, we have that

$$\begin{aligned}
\lim_{N \to \infty} \int_0^1 \int_0^1 f\left(p_1, \frac{z-\alpha p_1}{1-\alpha}\right) dH_N(p_1, z) &= \int_0^1 \int_0^1 f\left(p_1, \frac{z-\alpha p_1}{1-\alpha}\right) dH(p_1, z) \\
&= \int_0^1 f\left(p_1, \frac{1/2-\alpha p_1}{1-\alpha}\right) dp_1 \\
&= (1-\alpha)\, g(1/2),
\end{aligned}$$

where g is defined as in (6). Now

$$\lim_{N\to\infty} N\Pr\{R_1+R_2=N\} = \lim_{N\to\infty}\frac{1}{2}\frac{2N}{N_2+1}E_{\psi_N}f(p_1,p_2)$$
$$=\frac{1}{2}\frac{1-\alpha}{1-\alpha}g\left(\frac{1}{2}\right)=\frac{1}{2}g\left(\frac{1}{2}\right).$$

Thus to complete the proof it is only necessary to show that the distribution of Z_N converges weakly to a distribution concentrated on 1/2.

Let $\mathbf{P}_N$ be the random variable (p_{1N},p_{2N}) with density ψ_N. Then, (11) states that conditional on a parameter n which is uniformly distributed on $0, 1,\dots, N_1$, $\mathbf{P}_N$ is distributed as two independent beta distributions.

Thus

$$E(\mathbf{P}_N)=EE(\mathbf{P}_N|n)=E\left(\frac{n+1}{N_1+2},\frac{N-n+1}{N_2+2}\right)=\left(\frac{1}{2},\frac{1}{2}\right)$$

since $E(n)=N_1/2$.

Let $V(\mathbf{P}_N)$ be the variance–covariance matrix of $\mathbf{P}_N$. Then

$$V(\mathbf{P}_N)=EV(\mathbf{P}_N|n)+VE(\mathbf{P}_N|n)$$
$$=E\begin{bmatrix} V(p_{1,N}|n) & 0 \\ 0 & V(p_{2,N}|n)\end{bmatrix}$$
$$+\begin{bmatrix} \dfrac{1}{(N_1+2)^2} & -\dfrac{1}{N_1+2}\dfrac{1}{N_2+2} \\ -\dfrac{1}{N_1+2}\dfrac{1}{N_2+2} & \dfrac{1}{(N_2+2)^2}\end{bmatrix} V_N(n),$$

where $V_N(n)=((N_1+1)^2-1)/12$ is the variance of a random variable uniformly distributed on $(0,\dots,N_1)$.

Now,

$$V(p_{1,N}|n)=\frac{n+1}{N_1+2}\left(1-\frac{n+1}{N_1+2}\right)\frac{1}{N_1+3}\leqslant\frac{1}{4}\frac{1}{N_1+3}$$

and similarly

$$V(p_{2,N}|n)\leqslant\frac{1}{4}\frac{1}{N_2+3},$$

so that

$$\lim_{N\to\infty} V(\mathbf{P}_N) = \frac{1}{12}\begin{bmatrix} 1 & -\dfrac{\alpha}{1-\alpha} \\ -\dfrac{\alpha}{1-\alpha} & \left(\dfrac{\alpha}{1-\alpha}\right)^2 \end{bmatrix},$$

where $\alpha = \lim_{N\to\infty} N_1/2N$. Let $\boldsymbol{\alpha}' = (\alpha, 1-\alpha)$.

If $Z_N = \alpha p_{1,N} + (1-\alpha) p_{2,N} = \boldsymbol{\alpha}'\mathbf{P}_N$, then $E(Z_N) = \boldsymbol{\alpha}' E(\mathbf{P}_N) = 1/2$ and

$$\lim_{N\to\infty} V(Z_N) = \boldsymbol{\alpha}' \lim_{N\to\infty} V(\mathbf{P}_N)\,\boldsymbol{\alpha} = 0,$$

so $Z_N \to 1/2$ in quadratic mean and therefore in distribution. This completes the proof.

APPENDIX: A NOTE ON THE LITERATURE

Beck [2] made the first attempt of which we are aware to calculate the probability of having a decisive vote. Beck calculated $Q(N,p)$ for $p = 1/2$ and for p near 1/2 and observed the "knife-edged" nature of the result and concluded that the probability of casting a deciding vote "will only be significant ... if an individual assumes all other voters are totally indifferent between the two [alternatives]." Beck also considered a model in which there are two groups with different (but certain) probabilities of voting yes. His analysis of this case is, as he recognizes, tentative and incomplete. Margolis [5] extended Beck's analysis in somewhat the same way we did. He introduced subjectivity by supposing that to any voter, the fraction of the other voters voting yes is a random variable, y, with a probability distribution with density ϕ. If $2N$ people vote, then the subjective probability of a tie is $\int_{1/2-1/4N}^{1/2+1/4N} \phi(y)\,dy$, which is approximately equal to $(1/2N)\,\phi(1/2)$. This is, of course, the same result as we obtained in Proposition 1, where ϕ simply replaces f. It is, however, hard to understand how Margolis justifies his approach. The fraction of people voting yes is, for any finite election, a discrete random variable which can only take on $N+1$ values if N people vote. Margolis argues that ϕ is a normal distribution by appealing to the central limit theorem (or the normal approximation to the binomial). However, this is not legitimate in the present context. The distribution of the fraction of successes from a mixture of binomial distributions does not approach a normal distribution but instead approaches the mixing distribution itself. This fact, proved rigorously below, shows why Margolis' reasoning produces an answer in agreement with our Proposition 1.

PROPOSITION 4. *Let* $\Pr\{R = n\} = \int_0^1 \binom{N}{n} p^n (1-p)^{N-n}\, dF(p)$. *Then if* $N \to \infty$, $\Pr\{R/N \leqslant \tau\}$ *converges weakly to* F.

Proof. $\Pr\{R/N \leqslant \tau) = \int_0^1 \Pr\{R/N \leqslant \tau | p)\, dF(p)$. Let

$$\begin{aligned} I_\tau(p) &= 0 \qquad \text{if} \quad \tau < p \\ &= 1 \qquad \text{if} \quad \tau \geqslant p. \end{aligned}$$

The weak law of large numbers implies

$$\lim_{N\to\infty} \Pr\left\{\frac{R}{N} \leqslant \tau | p\right\} = I_\tau(p) \qquad \text{if} \quad \tau \neq p.$$

Since $\Pr\{R/N \leqslant \tau | p\} \leqslant 1$ and $\int_0^1 1 dF(p) = 1$, dominated convergence implies

$$\lim_{N\to\infty} \Pr\{R/N \leqslant \tau\} = \int_0^1 I_\tau(p)\, dF(p) = F(\tau)$$

if F is continuous at τ. Hence $\Pr(R/N \leqslant \tau)$ converges weakly to F.

REFERENCES

1. Y. BARZEL AND E. SILVERBERG, Is the act of voting rational? *Public Choice* **16** (1973), 51–58.
2. N. BECK, A note on the probability of a tied election, *Public Choice* **23** (1975), 75–79.
3. P. DIACONIS, Finite forms of de Finetti's theorem on exchangeability, *Synthese* **36** (1977), 271–281.
4. J. FEREJOHN AND M. FIORINA, The paradox of not voting: A decision theoretic analysis, *Amer. Pol. Sci. Rev.* **68** (1974), 525.
5. H. MARGOLIS, Probability of a tie election, *Public Choice* **31** (1977), 134–137.

JOURNAL OF ECONOMIC THEORY 25, 163 (1981)

Papers to Appear in Forthcoming Issues

The following articles are among those newly accepted for publication in the *Journal of Economic Theory*. They are listed alphabetically by first author.

TARADAS BANDYOPADHYAY, RAJAT DEB, AND PRASANTA K. PATTANAIK, The Structure of Coalitional Power under Probabilistic Group Decision Rules. Correspondent: Professor Taradas Bandyopadhyay, Department of Economics, Hull University, Hull, United Kingdom.

PIERRE BATTEAU, Revelation of Preferences for Public Goods: Aggregation by Abelian Operations. Correspondent: Professor Pierre Batteau, I.A.E., 29 Avenue R. Schuman, 13617 Aix en Provence, France.

JACQUES CRÉMER, A Simple Proof of Blackwell's "Comparison of Experiments" Theorem. Correspondent: Professor Jacques Crémer, Department of Economics, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia, PA 19104.

PETER C. FISHBURN, A Note on Linear Utility. Correspondent: Dr. Peter C. Fishburn, Bell Laboratories, Murray Hill, NJ 07974.

JOHN D. GEANAKOPLOS AND HERAKLIS M. POLEMARCHAKIS, We Can't Disagree Forever. Correspondent: Professor H. M. Polemarchakis, Department of Economics, Columbia University, New York, N.Y. 10027.

TERRY HEAPS, Location and the Comparative Statics of the Theory of Production. Correspondent: Professor Terry Heaps, Department of Economics, Simon Fraser University, Burnaby, B.C., V5A 1S6, Canada.

HARUO IMAI, CBD Hypothesis and Economies of Agglomeration: A Reexamination of Mills' Results. Correspondent: Professor Haruo Imai, Department of Economics, University of Southern California, Los Angeles, CA 90007.

LESTER G. TELSER, Voting and Paying for Public Goods: An Application of the Theory of the Core. Correspondent: Professor Lester G. Telser, Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637.

PATRICIA B. REAGAN AND MARTIN L. WEITZMAN, Asymmetries in Price and Quantity Adjustments by the Competitive Firm. Correspondent: Professor Martin L. Weitzman, Department of Economics, Room E52-353, Massachusetts, Institute of Technology, Cambridge, MA 02139.

Printed in Belgium

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved

How To Comply With The New Copyright Law

Libraries everywhere have found the easy way to fill photocopy requests legally and instantly, without the need to seek permissions, from this and over 3000 other key publications in business, science, humanities, and social science

Participation in the Copyright Clearance Center (CCC) assures you of legal photocopying at the moment of need. You can

Fill requests for multiple copies, interlibrary loan (beyond the CONTU guidelines), and reserve desk without fear of copyright infringement

Supply copies simply and easily from registered publications The CCC's flexible reporting system accepts photocopying reports and returns an itemized invoice *You need not keep any records*, our computer will do it for you

The Copyright Clearance Center is your one-stop place for on-the-spot clearance to photocopy for internal use You will never have to decline a photocopy request or wonder about compliance with the law for any publication registered with the CCC

For more information, just contact

Copyright Clearance Center
21 Congress Street
Salem, Massachusetts 01970
(617) 744-3350

a not-for-profit corporation

NAME ______________________ TITLE ______________________

ORGANIZATION ______________________

ADDRESS ______________________

CITY ______________________ STATE __________ ZIP __________

COUNTRY ______________________ TELEPHONE ______________________

JOURNAL OF ECONOMIC THEORY

Volume 25, Number 2, October 1981

Copyright © 1981 by Academic Press, Inc.
All Rights Reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc. (21 Congress Street, Salem, Massachusetts 01970), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre-1981 articles are the same as those shown for current articles.

Published bimonthly at 41 Tempelhof, Bruges, Belgium by
Academic Press, Inc., 111 Fifth Avenue, New York, N.Y. 10003.

1981: Volumes 24–25. Price $138.00 U.S.A.; $165.00 outside U.S.A.
1982: Volumes 26–28. Price: $192.00 U.S.A.; $229.50 outside U.S.A.
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publisher. All correspondence and subscription orders should be addressed to the office of the Publishers at 111 Fifth Avenue, New York, N.Y. 10003. Send notices of change of address to the office of the Publishers at least 6 to 8 weeks in advance. Please include both old and new addresses. Postmaster: Send changes of address to *Journal of Economic Theory*, 111 Fifth Avenue, New York, New York 10003.

Second class postage paid at Jamaica, N.Y. Air freight and mailing in the U.S.A. by Publications Expediting, Inc., 200 Meacham Avenue, Elmont, New York 11003.

Copyright © 1981 by Academic Press, Inc.

Printed in Bruges, Belgium, by the St. Catherine Press, Ltd.

JOURNAL OF ECONOMIC THEORY **25**, ii (1981)

Notice

We are pleased to announce that in 1982, *Journal of Economic Theory* will be expanded to three volumes of two issues each. During 1982, Volumes 26 through 28 will be published; the frequency of publication will be bimonthly with the first issue of the year appearing in February.

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

JOURNAL OF ECONOMIC THEORY 25, 165–195 (1981)

An Equilibrium Analysis of Search and Breach of Contract
II. A Non-Steady State Example

P. A. DIAMOND AND ERIC MASKIN

Department of Economics, MIT, Cambridge, Massachusetts 02139

Received May 8, 1979; revised June 12, 1980

1. INTRODUCTION

Consider a market where individuals meet pairwise and where each individual makes at most one trade (or is part of at most one project). The market for waterfront summer rentals is one example. Assume that search is costly and that there are fixed (and equal) numbers of potential landlords and renters. Suppose that tastes differ in that two renters may disagree about which of two houses is better, but assume that houses are ex ante identical in the sense that the distribution of evaluations is the same for each. For simplicity, assume only two possible evaluations: good and poor. In this paper, we analyze the equilibrium path over time of search and trade in such a market. We are particularly interested in its deviations from efficiency.

When the market opens, potential renters search for houses to rent, and landlords seek potential renters. Searchers meet according to a Poisson process. When a meeting results in a good match (a meeting where the renter's evaluation of the house is "good"), a rental is negotiated, a lease is signed, and both parties stop searching. When a poor match occurs, the parties also negotiate and sign a contract if neither already has a partner. However, they both search for a better deal if the expected benefits of further search exceed the costs. If neither partner finds a better deal, the two will ultimately carry out the negotiated rental. If one is successful[1], however, he can break the original lease and compensate his former partner for the loss borne. This better match might be with an individual who has not previously signed a lease. In that case, breaking the original lease is called a *single breach*. Alternatively, the new partner may already be matched, albeit poorly. Then, forming a new arrangement involves breaking both the old leases, a *double breach*. The process terminates when no one wishes to

---

[1] In our model the probability of simultaneous meetings is zero, so we can ignore the possibility that they both find better matches at the same time.

0022-0531/81/050165-31$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

continue searching, at which point all contracts remaining in existence are carried out.

We consider two distinct meeting processes or search technologies. In the first, the probability of a given renter's meeting a given landlord (assuming they are both searching) is independent of the number of other searchers. This process is designated the *quadratic case* since the rate of meetings rises with the square of the number of searchers. In the second meeting process, the probability of an individual's meeting someone at all is independent of the number of searchers. This is the *linear* technology.

With the quadratic technology, an equilibrium path[2] can assume several different patterns. What follows is one possibility. From the start of the search process until some time t_1, everyone who has not yet found a good match continues searching, and individuals make double and single breaches whenever they form a good match. From t_1 to t_2, those without good matches continue to search but now make only single breaches, since double breach is no longer worthwhile. At t_2, search ceases to be profitable for anyone already matched, and so only the partnerless continue. Finally, at time t_3, even they no longer find search worthwhile, and the process terminates. This sequence exhibits four different modes of behavior or configurations: A, where all individuals without good matches search, making both single and double breaches; B, where again those without good matches search but only single breaches occur; C, where only the partnerless search; and D, where there is no search. Since the order is A, B, C, D, the example is an ABCD sequence. We show below that the only other sequences possible are ACD, AD, BCD, CD, and D. Which one in fact occurs depends on the parameters of the economy and the starting point.

Besides describing equilibrium, the paper considers two efficiency questions. The first concerns the efficiency implications of small perturbations in the times t_i where the equilibrium path changes configurations. Returning to the above example and assuming that behavior is not changed before t_1, an increase in t_1, corresponding to a postponement of the cessation of double breaches, increases aggregate net output. In the neighborhood of t_1, a double breach alters the search environment in a way that others find valuable. The external economy implies, therefore, that, in equilibrium, Configuration A ends prematurely. Efficiency is also improved by increasing t_2; that is, aggregate output is increased if the time when those with poor matches stop searching is delayed. By contrast, perturbing t_3, the time when the partnerless stop searching, cannot improve efficiency. Because everyone searching at t_3 is identical (i.e., partnerless), each finds search to have zero marginal value at exactly the time when the social value of search (the sum of the individual values) is also zero.

---

[2] By "equilibrium path" we mean the trajectory that the numbers of searchers follow when, at each instant, every searcher maximizes his expected net gain, given the behavior of others.

Our second approach to efficiency is analysis of the optimal path. Here we derive the first order conditions of the times when configurations change. One interesting conclusion we draw is that the optimal path never involves Configuration B behavior (search by everyone without a good match but no double breach).

These results apply to the quadratic technology. With the linear technology, the perceived environment never changes for a typical searcher, and, therefore, behavior is uniform throughout the process. The equilibrium, moreover, is efficient.

This paper is concerned with much the same issues as was our previous study [1]. The principal distinction between the papers is that formerly we postulated the continuous arrival of new searchers and thus analyzed steady state behavior, whereas now we are concerned with the evolution of the search process.

After setting up the model (Section 2), the paper begins with consideration of the quadratic process: Section 3 examines the equilibrium path assuming full compensation for breach of contract; Section 4, the efficiency implications of small changes from the equilibrium path; Section 5, the efficient path; and Section 6, the equilibrium path assuming no compensation for breach. The linear process is presented in Section 7. We conclude with a brief summary and some remarks on the generality of the results.

2. THE MODEL[3]

We consider a model with two types of individuals.[4] Individuals are distinguished by type only in that each partnership (contract) requires exactly one partner of each type. Individuals search for a partner (of the opposite type) with whom to undertake a single project. If partners are well-matched, the project is worth $2X$. If not, output is $2X'$. We assume $X > X' > 0$. After partners have stopped searching—and only then—the project corresponding to their partnership is completed. Individuals are risk neutral and are able to make side payments with no bankruptcy constraints. Each individual can engage in at most one project and belong to at most one partnership.

Individuals can meet new potential partners only if they search, and the cost of search is a flow, c, per unit time. Under the quadratic technology, the probability that any two searchers (of opposite types) meet is a per unit time. Under the linear technology, a is the probability per unit time that a given searcher meets someone at all. We assume a is sufficiently small so that we can ignore the possibility that two partners who are both searching will simultaneously find new potential partners.[5] When two individuals meet, the probability of their matching poorly is p, with $1-p$, the probability they are a good match. All parameters are the same for individuals of both types, and so we shall refer to just one type.

---

[3] The model described below is essentially that of our earlier paper [1] and, therefore, is not discussed so fully as before.

[4] For example, buyers and sellers or lessors and lessees.

Let a partnerless individual be designated by "M," and let "N" refer to an individual with a poor contract. h_M denotes the number of M's, and h_N, the number of N's. For most of our analysis, we can disregard the number of individuals with good contracts, since they never search. When the search process begins there are $h_M(0)$ M's and no N's.

As in the Introduction, we classify search and breach behavior among four configurations. As we shall see, if two M's meet it will be in their interest to sign a contract regardless of the quality of match. However, N's, will breach only to form good contracts. Our interest in breaching behavior centers on whether N's will breach to form good contracts with both N's (double breaches) and M's or exclusively with M's (single breaches). Ceteris paribus, it is more advantageous, as we explain below, to form a contract with an M than with an N. Similarly, search is at least as profitable for an M as for an N. Thus there are three possibilities for search: it may be unprofitable for everyone, profitable for M's but not for N's, or profitable for both M's and N's. The last case subdivides in two breaching possibilities—either both single and double breaches are advantageous or only single breaches. The four possible behavior models are shown in Table 1. The effects of various meetings on the number of M's and N's are shown in Table 2. To interpret the first column of Table 2, we note that each of the h_M M's has a probability ah_M of meeting an M of the opposite type. p of these meetings result in poor matches, giving a flow of aph_M^2 new poor matches from meetings between M's. Each such meeting decreases the number of M's (of each type) by one and increases the number of N's by one. The equations determining the

TABLE 1

Behavioral Configurations

	Search by M's	Search by N's	Single breaches	Double breaches
A	yes	yes	yes	yes
B	yes	yes	yes	no
C	yes	no	irrelevant	
D	no	no	irrelevant	

---

[5] We have implicitly modeled contracting as instantaneous. Without instantaneous contracting, the assumption of no simultaneous meetings is an approximation.

TABLE 2

Numbers of Searchers, Quadratic Technology

Action	Poor match of M's	Good match of M's	Single breach	Double breach
Rate of flow	aph_M^2	$a(1-p)h_M^2$	$2a(1-p)h_M h_N$	$a(1-p)h_N^2$
Change in numbers of each type:				
without partners (h_M)	-1	-1	0	$+1$
in poor matches (h_N)	$+1$	0	-1	-2
in good matches	0	$+1$	$+1$	$+1$

number of searchers (see below) are calculated by multiplying the induced changes by the frequency of different types of meetings.

Under Configuration A, both M's and N's search, and any good match results in a contract's being signed. From Table 2 we can infer that the equations[6] are:

$$\begin{aligned} \dot h_M &= -ah_M^2 + a(1-p)\,h_N^2, \\ \dot h_N &= aph_M^2 - 2a(1-p)\,h_M h_N - 2a(1-p)\,h_N^2. \end{aligned} \tag{1}$$

The number of M's is diminished by matches between M's(ah_M^2) and enlarged by the individuals left partnerless when two N's make a good match ($a(1-p)\,h_N^2$). The number of N's is increased by poor matches between M's (aph$_M^2$) and diminished by good matches between N's and M's ($2a(1-p)\,h_M h_N$) and between N's ($2a(1-p)\,h_N^2$). Notice that

$$\begin{aligned} \dot h_M &\gtrless 0 \quad \text{as} \quad h_M \lessgtr h_N(1-p)^{1/2}, \\ \dot h_N &\gtrless 0 \quad \text{as} \quad h_M \gtrless h_N((1-p)^{1/2} + (1+p)^{1/2})(1-p)^{1/2}p^{-1} \end{aligned} \tag{2}$$

Under Configuration B, behavior is the same as under A except that double breach does not occur. The equations are, therefore, the same as (1) after deleting all terms involving double breaches:

$$\begin{aligned} \dot h_M &= -ah_M^2, \\ \dot h_N &= aph_M^2 - 2a(1-p)\,h_M h_N, \end{aligned} \tag{3}$$

$$\begin{aligned} \dot h_M &< 0, \\ \dot h_N &\gtrless 0 \quad \text{as} \quad h_M \gtrless 2h_N\left(\frac{1-p}{p}\right). \end{aligned} \tag{4}$$

---

[6] All differential equations are equations in the mean, to avoid stochastic components. Thus we are assuming that numbers are sufficiently large to realize the expected number of meetings.

Under Configuration C, only M's search. Therefore the equations are:

$$\dot{h}_M = -ah_M^2 < 0,$$
$$\dot{h}_N = aph_M^2 > 0. \tag{5}$$

Finally, under Configuration D, there is no search at all.

We define the positional values of being an M or an N, when the numbers of M's and N's are (h_M, h_N), as $V_M(h_M, h_N)$ or $V_N(h_M, h_N)$, respectively. An individual's positional value is the payoff that he can expect, given a correct forecast of the path of the economy and rational search and breach behavior. The equations of motion depend on the prevailing configuration and the numbers of M's and N's. The private decisions about search and breach that determine the configuration depend, in turn, on the future evolution of these equations of motion and the values of positions. Thus positional values are, indeed, functions solely of (h_M, h_N).

For positional value to be well-defined, we must describe how contracting works. Suppose that individuals i and j contemplate signing a contract which would yield them a combined positional value of $2V$. Let V^i and V^j be their current (i.e., pre-contract) positional values. Suppose further that i and j currently are in contracts which specify that they pay damages D^i and D^j, respectively, to their partners if they breach.[7] Then, we can define the surplus of the contemplated contract as $S = 2V - V^i - V^j - D^i - D^j$. We postulate that if i and j sign the contract, then they divide output and/or make side payments so as to share the surplus equally.[8] The individuals gain by signing the contract if and only if S is positive. Under our division rule, individuals i and j attain positional values $V^i + 1/2\,S$ and $V^j + 1/2\,S$, respectively, from the contract.

If, say, i breaches a contract with j, we assume that he pays j damages equal to $V^j - V_M$. (Recall that V_M is the positional value of being partnerless.) That is, j maintains the same expected payoff he had before the contract was breached. The damages are, therefore, compensatory. We focus on compensatory damage rules in this paper because (1) they constitute the basic principle for assigning damages under common law and (2) they are efficient in the limited sense that they ensure that a breach occurs if and only if there is an increase in the combined positional value of the principal affected parties[9] (the new partners and their original partners, if any). In Section 6 we consider equilibrium without damage payments. In our previous article we also considered liquidated damages, the damages that contracting parties themselves would choose.

---

[7] If i or j is currently partnerless, then $D^i = 0$ or $D^j = 0$, respectively.

[8] This division rule is also the Nash Bargaining Solution to the problem.

[9] For analysis of this effect of compensatory damages, see Mortensen [2].

In our simple two-quality model, two M's always wish to form a contract if they meet, since $V_N \geqslant V_M$ and $X > V_M$. Furthermore, with compensatory damages an N will breach his current contract only when he finds a good match, because only then can the surplus of his new contract be positive.

Notice that the incremental benefit that an M receives from any potential contract is larger than it would be were he an N, since he does not pay damages (which would diminish the surplus) and has a lower positional value. Therefore, the benefits of search are greater for M's than for N's. We need not consider, therefore, a behavior configuration in which N's but not M's search. Similarly, since an N would prefer to form a new contract with an M rather than another N, we can exclude configurations where double but not single breaches occur. We consider only equilibria where, at any moment, all individuals in the same position behave identically. Thus Configurations A, B, C, and D are, indeed, collectively exhaustive.

3. Positional Values, Trajectories and Boundaries in a Decentralized Market: Quadratic Technology

We assume that an individual maximizes positional value when deciding whether to search and breach. Positional value, however, depends on the future evolution of the economy and so must be determined by working backwards. We are interested in Nash equilibrium time paths. An equilibrium path specifies a behavior configuration at each instant of time and has the property that each individual finds the behavior prescribed for him optimal given the specified behavior of the others. To calculate the evolution of an equilibrium path, we derive and solve differential equations for positional values in each of the configurations. We then examine which configuration at any instant is consistent with equilibrium. When several configurations are all consistent at some instant—i.e., when there are multiple equilibrium paths—we select the configuration involving the most search and breach. So, for example, we select Configuration A over B, C, or D. We assume everyone is partnerless at the start,[10] i.e., h_N is zero, and h_M, arbitrary.

We begin the analysis with Configuration C. We refer to the set of (h_M, h_N) pairs where Configuration C behavior occurs on some equilibrium path as Region C.[11] Only M's search in Region C, and, therefore, along a Region C trajectory, h_M is steadily declining, while h_N is increasing at only p times the rate of h_M's decrease. Because M's make better partners than N's the gains from search monotonically decline for both M's and N's. Thus, a transition from Region C to either A or B is impossible. Once all N's stop searching, they will never wish to resume. Consequently, V_M^C, the positional value of an M in Region C, depends only on h_M. In Region C an M incurs search costs $c\Delta t$ in a small interval of time Δt and finds a partner with probability $ah_M\Delta t$. The value of a position at t equals the expected positional value at $t+\Delta t$, less search costs. Thus

---

[10] In the Appendix we drop this assumption.

[11] Regions A, B, D are similarly defined. Each point in $h_M - h_N$ space is, by our assumption of unique paths, associated with a unique region.

$$\begin{aligned} V_M^C(h_M(t)) = &-c\Delta t + ah_M\,\Delta t(pX' + (1-p)X) \\ &+ (1 - ah_M\,\Delta t)\,V_M^C(h_M(t+\Delta t)), \end{aligned} \tag{6}$$

where $pX' + (1-p)X$ is the expected output from a match. Rearranging terms, defining $\Pi = pX' + (1-p)X$, letting Δt tend to zero, and substituting for $\dot{h}_M$ using (5) yields

$$-(ah_M^2)\frac{dV_M(h_M)}{dh_M} = c - ah_M\Pi + ah_M V_M^C(h_M). \tag{7}$$

Equation (7) completes the first piece of the analysis: calculation of the change in positional value in Region C.

Because h_M declines steadily in Region C, it ultimately reaches h_M', where the gain from search for the next instant is zero. At this point the search cost equals expected gross gain:

$$c = ah_M'\Pi. \tag{8}$$

Thus, search ceases at h_M', and the line $h_M = h_M'$ serves as the transition boundary between Regions C and D.[12] To find V_M^C, therefore, we solve (7) with terminal condition $V_M^C(h_M') = 0$:

$$V_M^C = \Pi - \frac{c}{a}\frac{\ln h_M}{h_M} + \frac{G}{h_M}, \tag{9}$$

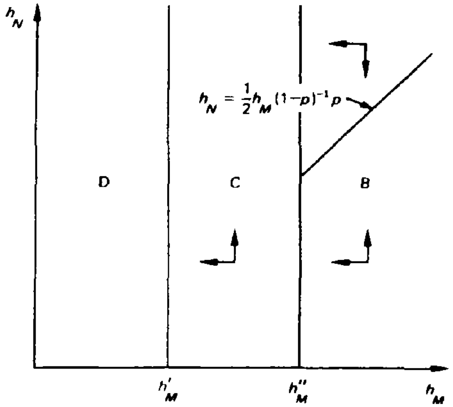
where $G = (c/a)(\ln c/a\Pi - 1)$.

Next consider possible transitions from B to C. In Configuration B N's make new deals only with M's. Thus, the transition boundary separating B from C falls at that critical number of M's h_M'', where an N finds search just barely profitable:

$$c = ah_M''(1-p)(X - X'), \tag{10}$$

---

[12] We refer to the locus of possible transitions as the transition boundary. Since (as we shall see) part of this locus may be in Region A or, alternatively, may not be reachable from an initial position on the h_M axis, the actual boundary of Region C is a subset of the transition boundary.



Arrows indicate direction of motion

FIG. 1. Directions of motion in Regions B, C, and D.

where $X - X'$ is one half the surplus of a contract between an M and an N when further search by N's is unprofitable. To understand this equation, note first that the value of the N position is X' since N's do no further search once Region C is reached. Thus, damages are $X' - V_M$, and the surplus from a new contract is $(2X - V_M - V_N - D) = (2X - 2V_N) = 2(X - X')$. Notice that $(1-p)(X - X') < \Pi$. Therefore $h''_M > h'_M$, and transition borders appear as in Fig. 1, where D borders only on C, and C only on B.

Next, consider positional values in Region B. Suppose that an M searches for time Δt beginning at time t. With probability $ah_M(1-p)\Delta t$, he will meet another M and form a good match, giving value X. If he meets an M and forms a poor match (probability $aph_M\Delta t$) his positional value becomes V_N. If he encounters an N with whom he makes a good match (probability $a(1-p)h_N\Delta t$) his positional value is V_M plus one-half the surplus of $2X - V_M - V_N - D = 2X - 2V_N$. Otherwise his positional value is $V_M(h_M(t+\Delta t), h_N(t+\Delta t))$. Thus, we have

$$
\begin{aligned}
&V^B_M(h_M(t), h_N(t)) \\
&\quad = -c\,\Delta t + a\,\Delta t\, h_M(1-p)X + a\,\Delta t\, h_M\, pV^B_N(h_M(t+\Delta t), h_N(t+\Delta t)) \\
&\quad\quad + a\,\Delta t\, h_N(1-p)(X - V_N(h_M(t+\Delta t), h_N(t+\Delta t)) \\
&\quad\quad + V_M(h_M(t+\Delta t), h_N(t+\Delta t))) \\
&\quad\quad + (1 - a\,\Delta t\, h_M - a\,\Delta t\, h_N(1-p))\, V_M(h_M(t+\Delta t), h_N(t+\Delta t)). \qquad (11)
\end{aligned}
$$

For an N, only a meeting with an M where a good match is made changes positional value. Positional value becomes $V_N + \frac{1}{2}(2X - 2V_N) = X$. Thus we have

$$\begin{aligned} V_N^B(h_M(t), h_N(t)) = {} & -c\,\Delta t + a\,\Delta t\, h_M(1-p)\,X + (1 - a\,\Delta t\, h_M(1-p)) \\ & \times V_N^B(h_M(t+\Delta t), h_N(t+\Delta t)). \end{aligned} \tag{12}$$

Taking limits we obtain the pair of differential equations

$$\begin{aligned} \frac{dV_M^B}{dt} &= c - a(1-p)(h_M + h_N)X - a(ph_M - (1-p)\,h_N)\,V_N^B + ah_M\,V_M^B, \\ \frac{dV_N^B}{dt} &= c - a(1-p)\,h_M X + a(1-p)\,h_M\,V_N^B. \end{aligned} \tag{13}$$

Using these equations, we can conclude that a transition from Region B to A is impossible. The surplus from a double breach is $S = X + V_M - 2V_N$. In Region B, S is negative since double breaches are unprofitable. Furthermore,

$$\frac{dS^B}{dt} = \frac{dV_M^B}{dt} - 2\,\frac{dV_N^B}{dt} = -c - a(X - V_N^B)((1-p)\,h_N + ph_M) + ah_M S < 0. \tag{14}$$

Therefore, the double breach surplus never becomes positive, and movement from B to A is ruled out.

Because transition from B to A is impossible and since $h_M'' > h_M'$, any trajectory crossing Region B must then move into Region C. (See Fig. 1.) Thus, once Region B is reached, N's never again contract with other N's and so V_N becomes a function of h_M alone. We can calculate positional values in B by solving the differential equation pair (13) with terminal conditions $V_M^B(h_M'', h_N) = V_M^C(h_M'')$ and $V_N^B(h_M'') = X'$. For $p \neq \frac{1}{2}$ we obtain

$$\begin{aligned} V_M^B(h_M, h_N) = {} & X - \frac{c(1-p)\ln h_M}{a(1-2p)\,h_M} + \frac{H}{h_M} \\ & - \frac{cp^2 h_M^{2p-2}}{a(1-p)(1-2p)^2\,(h_M'')^{2p-1}} \\ & + h_N \left[\frac{-c(1-p)\,h_M^{-2}}{ap(1-2p)} + \frac{ch_M^{p-2}}{ap(1-p)(h_M'')^p} \right. \\ & \left. + \frac{cph_M^{2p-3}}{a(1-p)(1-2p)(h_M'')^{2p-1}} \right]; \\ V_N^B(h_M) = {} & X + \frac{c}{aph_M} + \frac{J}{h_M^{1-p}}, \end{aligned} \tag{15}$$

where H and J are chosen so that the terminal conditions hold.

We now turn to an examination of Region A. For an M, the arrival of new opportunities in Region A follows the same rules as their arrival in Region B. Thus the differential equation for the change in value is the same. For an N, positional value falls more rapidly because of the added opportunity of double breaches. These occur at the rate $a(1-p)h_N$ and add $X - 2V^A_N + V^A_M$ to value when they occur. Thus the differential equations for values in Region A satisfy

$$
\begin{aligned}
\frac{dV^A_M}{dt} &= c - a(1-p)(h_M + h_N)X - a(ph_M - (1-p)h_N)V^A_N + ah_M V^A_M \\
\frac{dV^A_N}{dt} &= c - a(1-p)h_M(X - V^A_N) - a(1-p)h_N(X - 2V^A_N + V^A_M) \\
\frac{dS^A}{dt} &= \frac{dV^A_M}{dt} - 2\frac{dV^A_N}{dt} \qquad\qquad (16) \\
&= -c + a(1-p)(h_M + h_N)S - a((1-p)h_N + ph_M)(V^A_N - V^A_M) \\
\frac{d^2V^A_N}{dt^2} &= a(1-p)(h_M + h_N)c + a^2(1-p)ph_M(h_M + h_N)(V^A_N - V^A_M) > 0.
\end{aligned}
$$

We shall see that trajectories can move from Region A directly to any of the other three regions. However, there are two distinct patterns. For one set of parameter values, the equilibrium path moves from an initial position in A to B then C then D. For the remaining values, movement from A is directly

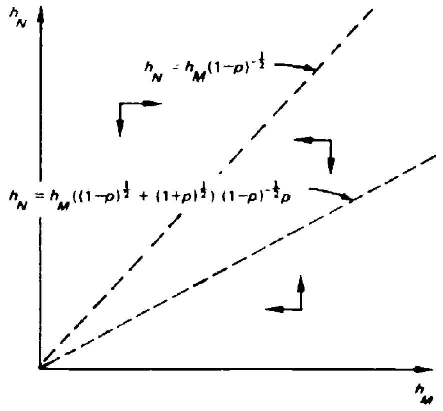

h_N
$h_N = h_M(1-p)^{-\frac{1}{2}}$
$h_N = h_M((1-p)^{\frac{1}{2}} + (1+p)^{\frac{1}{2}})(1-p)^{-\frac{1}{2}}p$
h_M

FIG. 2. Directions of motion in Region A.

to either C or D, and Region B does not exist. We first show that which of these two patterns applies depends on which region, A or B, contains the line $h_M = h''_M$. Then we consider, in turn, transitions from A to B, C, and D.

We observe from (2) that the direction of movement is as shown in Fig. 2. For an initial position on the h_M axis, the trajectory can never be to the left of the line $h_N = h_M(1-p)^{-1/2}$ while in Region A: movements across the line are not possible with Configuration A behavior. In the Appendix we briefly consider initial positions in this area. In the text we consider region boundaries only to the right of this line (although this restriction is often unstated).

Consider the point (h''_M, h_N) for some positive h_N. From Fig. 1 either this point is on the B–C border, or it is part of Region A. To check the former possibility, we need to evaluate the surplus from a double breach at (h''_M, h_N) using positional values in Region B. If the surplus is negative, (h''_M, h_N) is on the B–C border. If the surplus is positive, the point cannot lie in Region B. If (h''_M, h_N) is on the B–C border then, at this point, V_M equals $V^C_M(h''_M)$, V_N equals X', and we have

$$\begin{aligned} S &= 2X - 2V_N - 2D = 2X - 4V_N + 2V_M \\ &= 2\left(X - 2X' + \Pi - \frac{c}{a}\frac{\ln h''_M}{h''_M} + \frac{c}{ah''_M}\left(\ln\frac{c}{a\Pi} - 1\right)\right) \qquad (17) \\ &= 2(X - X')\left(1 + (1-p)\ln\frac{(1-p)(X-X')}{X' + (1-p)(X-X')}\right) \leqslant 0 \end{aligned}$$

Since (17) does not depend on h_N, either all the points (h''_M, h_N) lie on the

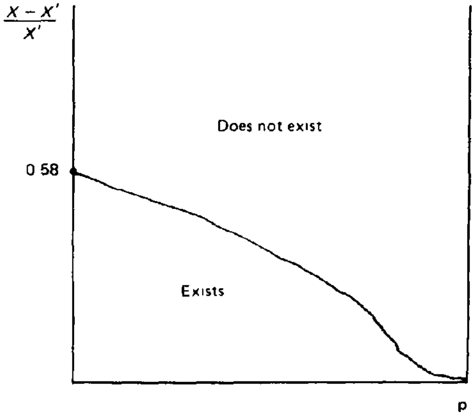


FIG. 3. Existence of Region B.

B–C border or they are all in Region A. Equation (17) is thus a necessary and sufficient condition for the existence of a B–C transition border.

Equation (17) indicates that for $X - X'$ sufficiently small relative to X' the surplus at $(h''_M h_N)$ is negative, and a B–C border exists. The smaller X relative to X', the less valuable is a good match and hence a double breach. Sufficiently small values of p also lead to a negative surplus. Decreasing p makes good matches easy to find, increasing the benefit to an N of remaining in the search market rather than taking advantage of a double breach opportunity. Therefore, small $X - X'$ (relative to X') or small p implies that Region B exists. These possibilities are shown in Fig. 3.

If (17) holds, we know that Region B exists and can integrate the value equations backward, ultimately reaching the A–B transition border. This border is defined as the locus where there is zero gain from a double breach:

$$X - 2V^B_N(h_M) + V^B_M(h_M, h_N) = 0. \tag{18}$$

Using Eqs. (15) and (16), we can write (18) as

$$\begin{aligned} h_N = h_M&[-(1-p)^2 + (1-2p)(h_M/h''_M)^p + p^2(h_M/h''_M)^{2p-1}]^{-1} \\ &\times [p(1-p)^2 \ln(h_M/h''_M) - 2(1-2p)(h_M/h''_M)^p \\ &+ p^3(1-2p)^{-1}((h_M/h''_M)^{2p-1} - 1) \\ &+ p(1-p)(1-2p)\ln((pX' + (1-p)X)/(1-p)(X-X')) \\ &+ (1-2p)(2-p)]. \end{aligned} \tag{19}$$

This locus is shown in Fig. 4. As h_M approaches h''_M, h_N increases without limit since $p^2 + (1-2p)$ equals $(1-p)^2$.

In Fig 4 we show the transition boundary. This suggests that Region A lies to the right, and B to the left, of locus (19). We have not confirmed that this is correct. To see the potential complication, consider the possibility that an equilibrium trajectory, when in Region A, might cross (19) more than once (see Fig. A1 and the discussion in the Appendix). If this were possible, only the last crossing would be a bonafide A to B transition, since we assumed the occurrence of the equilibrium path with the most breach. It would mean, furthermore, that Region A protrudes to the left of (19). We have not been able to rule out such multiple crossings. We can claim with accuracy, therefore, only that the A–B transition border is a subset of locus (19).

If S as defined by (17) is positive, the line $h_M = h''_M$ lies in Region A. A transition from B to A is impossible. Region B, if it exists, lies to the right of $h_M = h''_M$, and Regions C and D lie to the left. Therefore a positive S implies that no equilibrium path can go from Region B to another region, thus precluding the existence of B.

When B does not exist, we need to determine the A–C or A–D borders.

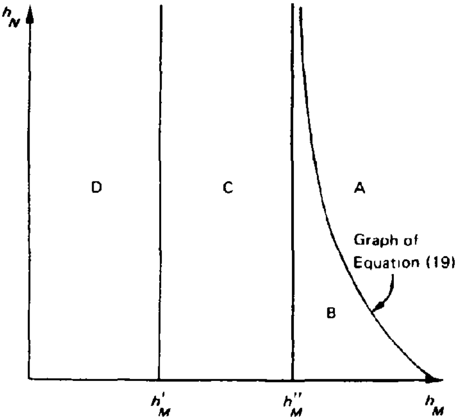


FIG. 4. Equilibrium regions when $X + V_M^C(h''_M) < 2X'$.

The A–C border is described by the curve showing indifference to continued search by N's:

$$c = a(1-p)\,h_M(X - X') + a(1-p)\,h_N(X - 2X' + V_M^C(h_M)). \tag{20}$$

The shape of the curve (20) depends on the sign of $X - 2X'$. Note that $c > a(1-p)\,h'_M(X - X')$. If $X < 2X'$, define h'''_M by $V_M^C(h'''_M) = 2X' - X$. Then $h'''_M > h'_M$, the A–C border lies to the right of the line $h_M = h'''_M$, and tends toward the line as h_N increases. This possibility is shown in Fig. 5. If $X > 2X'$, the A–C border reaches the line $h_M = h'_M$. In this case an A – D border exists and is given by

$$c = a(1-p)\,h_M(X - X') + a(1-p)\,h_N(X - 2X'). \tag{21}$$

This case is shown in Fig. 6. In Fig. 6 we have omitted indication of direction of movement in Region A. One may readily verify that the straight lines separating the areas of different direction of movement of Fig. 2 may bear any relation to the point of intersection of Regions A, C and D.

By analogy with the A–B border, we may inquire whether an equilibrium trajectory in Region A can cross the A–C ((20)) or A–D ((21)) transition borders more than once. Multiple crossings can be ruled out by the fact that an equilibrium trajectory is flatter than a 45 ° line ($dh_N/dh_M > -1$) (since the aggregate number of searchers is decreasing), whereas the transition are steeper than a 45° line: from implicit differentiation of (20) we have

$$\frac{dh_N}{dh_M} = -\frac{X - X' + h_N V_M'^C}{X - 2X' + V_M^C} < -1. \tag{22}$$

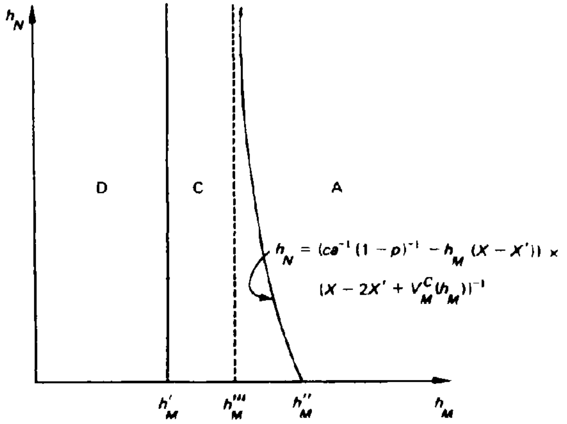

h_N
D
C
A
$h_N = (ca^{-1}(1-p)^{-1} - h_M(X - X')) \times (X - 2X' + V_M^C(h_M))^{-1}$
h_M'
h_M'''
h_M''
h_M

FIG. 5. Equilibrium regions when $X < 2X'$ and $X + V_M^C(h_M'') > 2X$.

The inequality follows since $V_M'^C > 0$ (see (9)) and $X' > V_M^C$ (since further search is not worthwhile). The same conclusion holds for the A–D border (21).

There is one remaining loose end to check. As one moves backward along trajectories in A, we must verify that search and double breach remain worthwhile. Moving backwards, both h_M and $(h_N + h_M)$ are increasing and so, therefore, is the return to search. Moving backward along a path, the surplus from a double breach cannot change sign (see (16)) since the other two terms are negative and dominate the term in S if S approaches zero.

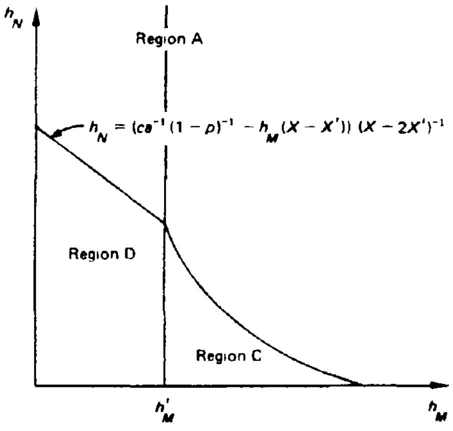


FIG. 6. Equilibrium regions when $X > 2X'$.

4. Inefficiency of Equilibrium

In Section 5, we describe efficient paths. In this section we examine the change in aggregate net output from perturbations of the equilibrium transition boundaries.[13] We show that the C–D boundary is efficient, but that shifts to the left of the B–C, A–C, and A–D borders (implying increased search) raise aggregate net output. The increase in breach resulting from a shift to the left in the A–B border also raises net output.

First consider the border between Regions C and D. If all M's search, the social gain per unit time is $ah_M^2(pX' + (1-p)X)$, while the social cost is ch_M.[14] Thus, the efficient C–D border, obtained by equating these expressions is the same as the competitive border. This coincidence may seem surprising, since an additional searcher creates an externality (an improvement in the positional value of other searchers) that does not seem to be captured by compensatory damages. The coincidence however is an artifice of the model's symmetry. The social gain from search is the sum of the individual gains. Since, under Configuration C, all searchers are identical, the social gain becomes zero precisely at the point where any individual gain vanishes. Thus the social and private incentives for search are the same.[15]

Next consider a slight shift to the left of the equilibrium B–C border. If the N's cease searching when $h_M = h_M''$, aggregate net output from those still in the market is $h_N X' + h_M'' V_M^C(h_M'')$. Suppose that N's continue to search an instant longer. The presence of N's following Configuration B behavior does not alter the time path of h_M, since single breaches do not affect the number of M's (Eqs. (3) and (5) are the same). Thus after the N's cease searching, the trajectory is the same as in Section 3. The cost of the N's additional search is ch_N per unit time. The additional output per unit time is the aggregate surplus from the resulting good matches, $2a(1-p)\, h_M h_N(X - X')$. The N's receive only half this surplus. Thus the private incentive to search is smaller than the social gain. A shift to the left of the B–C border (prolonging search by N's) raises net aggregate output since $c < 2a(1-p)\, h_M''(X - X')$.

We next consider perturbation of the equilibrium A–D border (which exists when $X > 2X'$). The border is the locus of points where N's are just willing to search (given that they receive half the surplus from both single and double breaches) and where M's are willing to search only if N's do so. M's find search worthwhile at the A–D border only because they receive part of the surplus from single breaches.[16] Again search ceases too soon—a shift to the left of the A–D border raises aggregate net output. The net gain from continued search is the surplus from matches between *M*'s, from single breaches, and from double breaches, minus the search costs:

---

[13] In doing these perturbations, we assume they do not affect earlier behavior before the perturbed boundary is reached.

[14] Aggregate output and search costs are twice these figures, but we continue to focus on one of the two types that make a pair.

[15] If individuals differed in search cost, all but the searchers with lowest cost would stop search too soon.

$$\begin{aligned} &ah_M^2((1-p)X+pX')+2a(1-p)\,h_M h_N(X-X') \\ &\quad + a(1-p)\,h_N^2(X-2X')-c(h_M+h_N). \end{aligned} \tag{23}$$

At the equilibrium border, (21), indifference of the *N*'s to continued search implies a net social gain from continued search of

$$h_M(ah_M((1-p)X+pX')+a(1-p)\,h_N(X-X')-c) \tag{24}$$

But from the indifference of *N*'s to search ((21), again) (24) becomes

$$ah_M(h_M+(1-p)\,h_N)\,X', \tag{25}$$

which is positive. To obtain the efficient A–D border, (23) is set equal to zero.

Consideration of the A–C border introduces a new element, not present in discussions of the other borders: the effect of double breach on the search environment as a result of changing the number of *M*'s. The external effect is irrelevant at the A–D border because all search ceases there. Let V^* be the aggregate value of continued search. Once Region C is reached we have V^* equal to $h_M V_M(h_M)$. Since V_M^C is increasing in h_M, a double breach at the A–C border generates an external economy. Thus both single and double breaches have social values that differ from their private values to *N*'s. The increase in aggregate value from continued search by *N*'s of both types is the full value of single and double breaches plus the increased value of the search process for *M*'s:

$$\begin{aligned} &2a(1-p)\,h_M h_N(X-X')+a(1-p)\,h_N^2(X-2X'+V_M^C(h_M)) \\ &\quad + a(1-p)\,h_N^2 h_M V_M'^C(h_M)-ch_N. \end{aligned} \tag{26}$$

Using (20), (26) becomes

$$a(1-p)\,h_M h_N(X-X')=a(1-p)\,h_N^2 h_M V_M'^C. \tag{27}$$

which is positive.

Thus a leftward shift in the A–C border, resulting in prolonged search by

---

[16] *M*'s find search at least as profitable as *N*'s. Thus when *N*'s are indifferent to search, as at the A–D border, *M*'s strictly prefer to continue searching.

N's, is desirable in part because of the externalities from double breaches. For the efficient border we set (26) equal to zero.

We turn, finally, to the A–B border. A slight shift of the equilibrium border induces no change in search but affects breaching behavior. We show that the continuation of double breaches beyond the A–B border is worthwhile, assuming the rest of the equilibrium process is unchanged. At the A–B border, a double breach yields no private gain; nevertheless, there is a social gain. The value of net output of continued search in Region B is

$$V^* = h_M V_M^B(h_M, h_N) + h_N V_N^B(h_M). \tag{28}$$

A double breach creates a good match, adds one M and subtract two N's. The impact of these changes on aggregate value is

$$\begin{aligned} \Delta V^* &= X + \frac{\partial}{\partial h_M}(h_M V_M^B + h_N V_N^B) - 2\frac{\partial}{\partial h_N}(h_M V_M^B + h_N V_N^B) \\ &= \left(\frac{c}{a}\right)\left(\frac{p(2-p)}{1-p} + 2p\frac{h_N}{h_M}\right)\left(\frac{(h_M/h_M'')^{2p-1} - 1}{h_M(2p-1)}\right) \\ &\quad + \left(\frac{c}{a}\right)\left(\frac{1}{1-p} + \frac{h_N}{h_M}\right)\left(\frac{1}{h_M}\right) > 0, \end{aligned} \tag{29}$$

since $h_M > h_M''$ on the A–B border. (We have used (15) to calculate this derivative.)

To understand the externalities created by breach, we can examine the impact of a double breach on individuals other than the four principal parties (the two breachers and their partners). An M gains $X - V_N$, whereas an N gains $X - 2V_N + V_M$ from a good match with an N. If the double breach does not occur, the principals remains N's. If breach occurs, an M gains $X - V_M$ and an N, $X - V_N$, from a good match with the principal party left partnerless, while an M gains $V_N - V_M$ from a poor match with this party. Neither M nor N gains anything from meeting the breacher (who is now well matched). Thus the sign of an M's net gain from double breach is the same as that of

$$\begin{aligned} &p(V_N - V_M) + (1-p)(X - V_M - 2(X - V_N)) \\ &\quad = p(V_N - V_M) + (1-p)(2V_N - X - V_M); \end{aligned} \tag{30}$$

whereas the sign of the N's gain is the same as that of

$$(1-p)(X - V_N - 2(X - 2V_N + V_M)). \tag{31}$$

At the A–B border, $X - 2V_N + V_M$ is zero. Hence both (30) and (31) are positive there. Thus, when the principal parties are themselves indifferent about carrying out a double breach, the overall externality induced by such a breach is positive.

5. THE EFFICIENT PATH

Since the C–D border is efficient, we can straightforwardly derive the efficient A–C and A–D borders from the perturbation analysis above. To complete the analysis, we verify that an efficient path can never cross from Region C to either A or B and that it never entails Configuration B behavior.

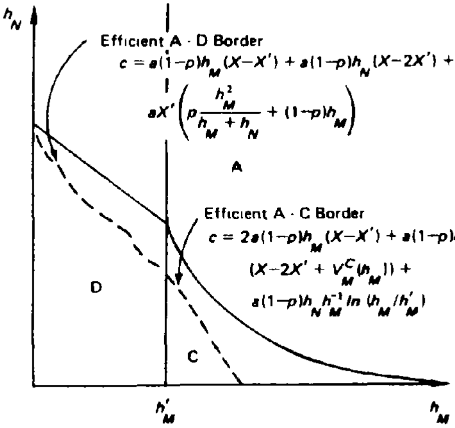
To describe the efficient A–C and A–D borders, we equate values of the perturbations of these borders with zero. Thus setting (23) and (26) equal to zero yields the border equations. In Figs 7, 8, and 9 we compare the efficient and equilibrium borders.

When $X > 2X'$ an A–D border exists and is obtained by setting (23) equal to zero:

$$a(1-p)X(h_M+h_N)^2 - 2a(1-p)h_N(h_M+h_N)X' + aph_M^2X' = c(h_M+h_N) \tag{32}$$

or

$$c = a(1-p)h_M(X-X') + a(1-p)h_N(X-2X') + aX'\left(p\frac{h_M^2}{h_M+h_N} + (1-p)h_M\right). \tag{33}$$



h_N
Efficient A · D Border
$c = a(1-p)h_M(X-X') + a(1-p)h_N(X-2X') +$
$aX'\left(p\frac{h_M^2}{h_M+h_N} + (1-p)h_M\right)$
A
Efficient A · C Border
$c = 2a(1-p)h_M(X-X') + a(1-p)h_N \times$
$(X-2X' + V_M^C(h_M)) +$
$a(1-p)h_N h_M^{-1} \ln(h_M/h'_M)$
D
C
h'_M
h_M

FIG. 7. Efficient and equilibrium borders when $X > 2X'$.

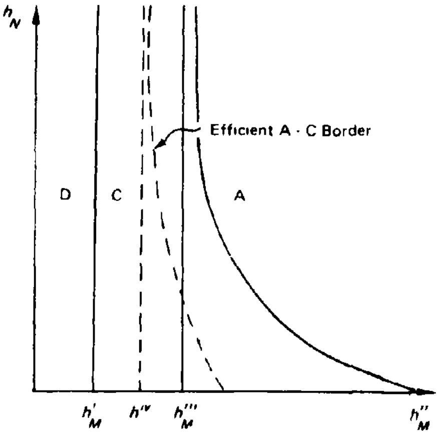


FIG. 8. Efficient and equilibrium borders when $X < 2X'$ and $X + V^C_M(h''_M) > 2X'$.

The equilibrium equation, (21), differs from (33); it does not contain the last term on the right. Thus for any value of h_N, there is a smaller value of h_M for the efficiency border than for the equilibrium border. For the A–C border, setting (26) equal to zero gives

$$\begin{aligned} c = {} & 2a(1-p)\,h_M(X-X') + a(1-p)\,h_N(X-2X'+V^C_M(h_M)) \\ & + c(1-p)\,h_N h_M^{-1} \ln(h_M/h'_M). \end{aligned} \tag{34}$$

When $X > 2X'$, the coefficients for h_M and h_N are both positive and, the right hand side of (34) exceeds that of (20), the equilibrium border equation. Thus the efficient border lies to the left of the equilibrium border. This relation and the A–D border are shown in Fig. 7.

In Fig. 8 we illustrate the case where $X - 2X' + V^C_M(h''_M)$ is positive and $X < 2X'$. Substituting for V^C_M (from (9)) one sees that the coefficient of h_N in (34) vanishes at h^{iv}_M satisfying

$$\frac{h^{iv}_M}{h'_M} = \frac{pX' + (1-p)X}{2X' - X}. \tag{35}$$

Thus the efficient border is asymptotic to this line.

The remaining case to consider is where the market equilibrium has a Region B. In this case the efficient A–C border, like its equilibrium counterpart, is asymptotic to the line $h_M = h^{iv}_M$. This is shown in Fig. 9.

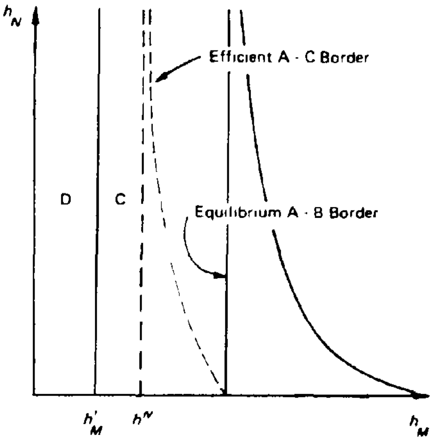


FIG. 9. Efficient and equilibrium borders when $X + V'_M(h''_M) < 2X'$.

As the figures show, we have too little search in equilibrium unless $h_M(0)$ is in Region C or Region D.[17]

This discussion of efficient borders implicitly assumed that N's never resume search after having stopped, i.e., that the efficient path does not enter Regions A or B from Region C. We now prove that such transitions are impossible. In Region C, the dynamic programming value equation for aggregate net output is

$$\begin{aligned} \dot{V}^* &= \frac{\partial V^*}{\partial h_M}\dot{h}_M + \frac{\partial V^*}{\partial h_N}\dot{h}_N = -ah_M^2\frac{\partial V^*}{\partial h_M} + aph_M^2\frac{\partial V^*}{\partial h_N} \\ &= ch_M - a(1-p)h_M^2 X, \end{aligned} \tag{36}$$

where $V^*(h_M, h_N)$ is aggregate value from those not in good matches. In addition, we know that, in Region C, the marginal social value of N is constant, since N's simply accumulate:

$$\frac{d}{dt}\left(\frac{\partial V^*}{\partial h_N}\right) = 0. \tag{37}$$

From these two equations we can contradict the rise in the value of search by N's which would necessarily accompany a transition from C to either B or A. Search by an N at time t has net value $2a(1-p)\,h_M(t)(X - \partial V^*/\partial h_N) - c$. In region C, $h_M(t)$ is decreasing and all other terms are constant. Furthermore $\partial V^*/\partial h_N < X$ because an N can, at best, make a good match. Thus the value of search declines, and so a C to B transition is impossible.

---

[17] We have assumed throughout that all N's stop searching at the same time. This assumption is justified since the social value of search by an N decreases when another N stops searching. That is, just to the right of the hypothesized borders, $\partial V^*/\partial h_N$ exceeds X', and, when these values are equal, it is socially worthwhile for all N's to stop searching simultaneously.

To consider a move from C to A, we must consider search by all the available N's. The aggregate return to search by N's is

$$2a(1-p)\,h_M h_N\left(X - \frac{\partial V^*}{\partial h_N}\right) + 2a(1-p)\,h_N^2\left(X - 2\frac{\partial V^*}{\partial h_N} + \frac{\partial V^*}{\partial h_M}\right) - ch_N.$$

Using (36) to eliminate $\partial V^*/\partial h_M$, we derive the return to search per N,

$$\begin{aligned}
&2a(1-p)\,h_M\left(X - \frac{\partial V^*}{\partial h_N}\right) + 2a(1-p)\,h_N\left(X - 2\frac{\partial V^*}{\partial h_N} + (ah_M)^{-1}\right.\\
&\qquad \left.\times\left(aph_M\frac{\partial V^*}{\partial h_N} + a(1-p)\,h_M X - c\right)\right) - ch_N\\
&\qquad = 2a(1-p)\,h_M\left(X - \frac{\partial V^*}{\partial h_N}\right) + 2a(1-p)\,h_N(2-p)\left(X - \frac{\partial V^*}{\partial h_N}\right)\\
&\qquad - c(2(1-p)\,h_N h_M^{-1} + h_N).
\end{aligned} \tag{38}$$

Differentiating with respect to time (and using (37)) we have

$$\begin{aligned}
&2a(1-p)\left(X - \frac{\partial V^*}{\partial h_N}\right)(\dot{h}_M + (2-p)\,\dot{h}_N)\\
&\qquad - c(2(1-p)((h_M\dot{h}_N - h_N\dot{h}_M)/h_M^2) + \dot{h}_N)\\
&\qquad = 2a^2(1-p)\left(X - \frac{\partial V^*}{\partial h_N}\right)(-h_M^2 + (2-p)\,ph_M^2)\\
&\qquad - ac(2(1-p)(ph_M + h_N) + ph_M^2) < 0,
\end{aligned} \tag{39}$$

since $(2-p)\,p < 1$. Thus the return to search by N's decreases per N while in Region C and, therefore, can never become positive.

We now turn to the proposition that the efficient trajectory never involves Configuration B behavior. We show that there is a higher net output flow from either Configuration A or C behavior if there is a border where the efficient trajectory leaves Region B[18]. This contradicts the possibility of Configuration B behavior on the efficient path. First consider the B–C border. The excess of aggregate net output flow that accrues from Configuration B behavior above that yielded by Configuration C is the additional output from single breaches less the search cost of N's,

---

[18] Nor can it stay indefinitely in Region B.

$$2a(1-p)\,h_M h_N\left(X-\frac{\partial V^*}{\partial h_N}\right)-ch_N. \tag{40}$$

If a B–C efficient border exists, it is the locus where (40) equals zero. Since N's search no further, $\partial V^*/\partial h_N = X'$ at the border. Therefore the border is defined by the equation $h_M = \hat{h}_M$, where

$$\hat{h}_M=\frac{c}{2a(1-p)(X-X')}. \tag{41}$$

The difference between the Congifuration A and B rates of aggregate net output is the social gain from double breaches:

$$a(1-p)\,h_N^2\left(X-2\frac{\partial V^*}{\partial h_N}+\frac{\partial V^*}{\partial h_M}\right). \tag{42}$$

At the B–C efficient border, $\partial V^*/\partial h_N = X'$ and $\partial V^*/\partial h_M = V_M^C(\hat{h}_M) + \hat{h}_M(dV_M^C(\hat{h}_M)/dh_M)$ where V_M^C is given by (9). Thus, at the border,

$$\begin{aligned} X-2\frac{\partial V^*}{\partial h_N}+\frac{\partial V^*}{\partial h_M} &= X-2X'+\Pi-\frac{c}{a\hat{h}_M}\\ &= X-2X'+\Pi-2(1-p)(X-X')\\ &= p(X-X')>0. \end{aligned} \tag{43}$$

Formula (42) is therefore positive, and so an efficient B–C border does not exist. That is, when the social return to search by N's is close to zero and that of M's is positive, double breach is socially worthwhile.

An efficient B–D border is similarly ruled out by the positive social value of double breaches. At a B–D border, $\partial V^*/\partial h_M = 0$, $\partial V^*/\partial h_N = X'$. and the gain from double breach, (42), is

$$a(1-p)\,h_N^2(X-2X). \tag{44}$$

At this border the additional net output from search is just zero. That is.

$$ah_M^2\Pi+2a(1-p)\,h_M h_N(X-X')=c(h_M+h_N). \tag{45}$$

Furthermore, search by N's must be socially worthwile at the border. That is. Configuration B must be more efficient than C. implying $h_M \geqslant \hat{h}_M$ or

$$2a(1-p)\,h_M(X-X')\geqslant c. \tag{46}$$

Subtracting h_N times (46) from (45), we obtain

$$ah_M^2 \Pi \leqslant ch_M. \tag{47}$$

From the condition $h_M \geqslant \hat{h}_M$, (47) becomes

$$c \geqslant a\hat{h}_M \Pi = \frac{c\Pi}{2(1-p)(X-X')}$$

or

$$2(1-p)(X-X') \geqslant (1-p)X + X'. \tag{48}$$

This last expression implies $X > 2X'$, implying a positive gain from double breaches ((44)), which is a contradiction.

The last remaining possibility is a transition from B to A. At the A–B border, the surplus from double breaches must be zero; hence

$$X - 2\frac{\partial V^*}{\partial h_N} + \frac{\partial V^*}{\partial h_M} = 0. \tag{49}$$

Under Configuration B, the dynamic programming value equation is

$$\begin{aligned} \frac{d}{dt} V^*(h_M(t), h_N(t)) &= \frac{\partial V^*}{\partial h_M}\dot{h}_M + \frac{\partial V^*}{\partial h_N}\dot{h}_N \\ &= -ah_M \frac{\partial V^*}{\partial h_M} + (aph_M^2 - 2a(1-p)\,h_M h_N)\frac{\partial V^*}{\partial h_N} \\ &= c(h_M + h_N) - a(1-p)(h_M^2 + 2h_M h_N)\,X. \end{aligned} \tag{50}$$

Solving for $\partial V^*/\partial h_N$ using (49) and (50) we obtain

$$\begin{aligned} -(a(2-p)\,h_M^2 + 2a(1-p)\,h_M h_N)\frac{\partial V^*}{\partial h_N} = c(h_M + h_N) - a((2-p)\,h_M^2 \\ + 2h_M h_N(1-p)X, \end{aligned} \tag{51}$$

or

$$X - \frac{\partial V^*}{\partial h_N} = \frac{c(h_M + h_N)}{a((2-p)\,h_M^2 + 2(1-p)\,h_M h_N)} < \frac{c}{2a(1-p)\,h_M}. \tag{52}$$

If Configuration B is at least as efficient as C, however, then

$$X - \frac{\partial V^*}{\partial h_N} \geqslant \frac{c}{2a(1-p)} \tag{53}$$

(i.e., non-negative value of search by N's). Since inequalities (52) and (53) are mutually contradictory, we conclude that there is no efficient Region B.

6. Equilibrium without Damages

We have considered equilibrium assuming an idealization of the common law's provision of compensatory damages. Often individuals do not avail themselves of these damages. One way of modeling this behavior is to assume that individuals do not sign contracts unless they stop searching. Instead, we assume that after an individual has found a poor match, he can fallback on that match when further search is unprofitable, provided the fallback partner is still available. We assume that only one fallback contact is preserved, and that individuals do not replace an earlier fallback with a later one as long as the earlier one is available.[19] This behavior is captured by the model in Sections 2 and 3 if damages are always set at zero. We assume that after two individuals make a poor match, the decision to stop searching and to complete the project is jointly made. If the search decision were not joint, each partner would find search individually profitable, assuming his fallback partner did not search, at the point where search becomes jointly unprofitable.

The C–D border is the same with or without damages since only M's are involved. Thus $V_M(h_M)$ is the same in Region C as previously. Region B does not exist, because a double breach for a good match is always profitable with zero damages. Thus we are interested in A–C and A–D transitions. We shall see that the absence of damages lowers the incentive to search. Damage payments come out of surplus before its division. Thus a new partner effectively pays half the damages to one's old partner. This monopoly power over new partners serves as a further incentive to search for the original partners when damages are positive.[20]

In Configuration A search by a pair of N's for additional time Δt costs each $c\,\Delta t$. Each has probability $a(1-p)\,h_M\,\Delta t$ of forming a good match with an M, yielding a gain to the pair of one-half the surplus, $\frac{1}{2}(2X - V_M - V_N)$, less the damages suffered (but not paid) of $V_N - V_M$. Each partner has probability $a(1-p)\,h_N\,\Delta t$ of a double breach with surplus $\frac{1}{2}(2X - 2V_N)$, and the same damages, $V_N - V_M$. When the pair is just willing to search, V_N equals X', and we have the A–C border equation

$$c = a(1-p)\,h_M(X + \tfrac{1}{2}V_M - \tfrac{3}{2}X') + a(1-p)\,h_N(X - 2X' + V_M). \qquad (54)$$

---

[19] These two assumptions are discussed below.

[20] This theme is explored in our previous paper, which also examines the damages which partners would choose to set (liquidated damages).

Because new partners do not share in damage payments, Eq. (54) differs from the A–C border with compensatory damages, (18), by $-a(1-p)h_M\frac{1}{2}(X'-V_M)$. (Symmetry implies no change in the gain from double breaches.) Thus there is less incentive to search without damages and the A–C border lies to the right of its position with damages. We can derive the A–D border from the A–C boder by setting V_M equal to zero. Again, there is less search than with compensatory damages.

The behavioral assumptions in this section have permitted a simple modification of the basic model. This modification illustrates the search incentive inherent in damages. With a further modification, we can illustrate the role of compensatory damages in joint maximization by partners. We have assumed so far that there were no single breaches for the sake of replacing one poor match with another. Yet individuals have an incentive to do so. When two M's form a poor match, they plan to evenly divide the output, $2X'$, should they carry out the match. If one of these N's meets a new M just as he is about to stop searching, the N can gain from breach even if his new partner is a poor match. Forming a new partnership, the N receives[21] $X' + \frac{1}{2}(X' - V_M)$ while the M receives $V_M + \frac{1}{2}(X' - V_M)$. The breach reduces the positional value of the previous partner from X' to V_M. Thus the aggregate positional value of the original partners has declined by one-half the compensatory damages that are not being paid. If the original partners have no way to control this inefficient breaching behavior, they will find search less profitable. The first order condition for the end of search will differ from that in the basic model, (18), by $-\frac{1}{2}ah_M(X' - V_M)$ rather than the factor $-\frac{1}{2}a(1-p)\,h_M(X' - V_M)$ in the first modification.

It is artificial to assume that individuals keep track of only one fallback partner. So too, in the basic model it is artificial to assume that individuals do not keep track of potential partners they have met with whom they do not form partnerships. Introducing a more complicated information structure would be interesting but would add considerably to the difficulty of analyzing the model.

7. LINEAR TECHNOLOGY

When the density of potential trading partners is low, the quadratic technology may be a reasonable approximation. However, when the density is high or the information about location is good, a searcher's problem is less one of finding a potential partner than of finding one who yields a high surplus. Such a situation can be approximated by assuming a constant probability of meeting someone at all, independent of the numbers of potential partners (although constancy is improbable if the numbers of searchers are small). Analysis is quite different from that above since the market possibilities do not alter as time (and the numbers of searchers) changes.[22] Thus the entire $h_M - h_N$ space is characterized by a single configuration, A, C, or D, depending on parameters.[23] That is, individuals search until they find a good match, or search until their first match, or do not search at all, What is more, the equilibrium path is efficient.

---

[21] Just as N's are due to stop searching, $V_N = X'$. The surplus from breach, $V_N - V_M$ is positive everywhere, however, not just at the A–C border.

There are three possibilities:

$c > a\Pi$	search is not worthwhile (Region D),
$a\Pi \geqslant c \geqslant a(1-p)(X-X')$	search is worthwhile for M's but not N's (Region C),
$c < a(1-p)(X-X')$	search is worthwhile for N's, implying that no bad matches are made; $V_N = V_M$ (Region A).

With this technology and behavior, positional values are independent of the numbers of searchers, giving

$$V_M^D = 0,$$

$$V_M^C = \Pi - \frac{c}{a},$$

$$V_M^A = V_N^A = X - \frac{c}{a(1-p)}.$$

Thus aggregate net output, V^*, is linear in h implying that the competitive process is efficient.[24]

8. BRIEF SUMMARY

We have studied an allocation mechanism in which a searcher's meeting opportunities arrive according to a Poisson process. In Configuration A under the quadratic technology, for example, poor opportunities arrive at the rate $aph_M(t)$ and good ones at $a(1-p)(h_M(t)+h_N(t))$. The values of these opportunities are determined endogenously; they depend on the evolution of the allocation process. The first part of the paper examines equilibrium evolutions: time paths where search and breach decisions are individually optimal.

---

[22] This result would change if those with poor contracts searched and the model were changed so that poor contracts were sometime carried out.

[23] If search continues until a good match is made, a poor match is of no additional value over being partnerless. Thus the damages are zero, and double breaches are always profitable, implying that Configuration B behavior does not occur.

[24] The efficiency of the competitive process under the linear technology is not robust to generalizations of the model. (See our earlier paper.)

An equilibrium time path consists of a sequence of behavior configurations determined by the parameters of the search technology (a, c), of tastes (p, X, X') and of initial position $(h_M(0))$. Section 3 enumerates all possible sequences. It demonstrates for example that the only equilibrium paths involving all four behavior Configurations are paths beginning in Region A and proceeding in turn to B, C, and D.

The heart of the paper is the demonstration that under the quadratic technology, search and breach give rise to externalities that generally cause inefficiency in the market process. Search by an individual creates a positive economy for other searchers. Because this externality is uncompensated, in equilibrium, N's stop searching too soon for efficiency. Double breach also creates external economies; it alters search environment by replacing two N's by an M on each side of the market. Since (at a point where double breach is individually just worthwhile), a searcher prefers the probability of meeting an M to twice that probability of meeting an N, such a replacement is a positive externality. Therefore, equilibrium paths entail too little breach; i.e., the transition from Region A to B occurs too soon for efficiency.

Given the simplicity of our model, it is natural to question whether the results are robust. We believe that the conclusion that double breach induces a positive externality when it is marginally worthwhile for individuals is quite general; the same result obtains in a variety of other models we have explored, including one with a continuum of qualities. Moreover, it is not a result peculiar to the quadratic technology. It applies, for example, to linear models that are sufficiently general so that poor contracts are sometimes carried out. (This illustrates that our conclusion that equilibrium under the linear technology is efficient does not generalize). Unfortunately, the reasons underlying the result are too complicated to go into here; they will be the subject of a forthcoming paper.

Whether adding an additional searcher creates a positive or negative externality depends on the value of the added searcher as a potential partner and on the search technology. For all technologies that we have examined—quadratic linear, and everything in between—adding a searcher s has two potential effects: (1) to raise to total probability of finding a potential partner and (2) to reduce the probability of finding a potential partner other than s. In the quadratic technology, only effect (1) is present, and so additional searchers always induce positive externalities, no matter how complicated the model otherwise is. In the linear technology, only effect (2) operates, and so the sign of the externality depends on s's value as a partner: positive if s is relatively valuable (an M, in the model of this paper) and negative if not (an N in the present model). (Externalities of search in the linear technology are irrelevant to efficiency in the linear model of this paper but relevant in more general linear models.) In models intermediate between quadratic and linear, both effects pertain, and so again a searcher's potential value will determine the sign of the externality.

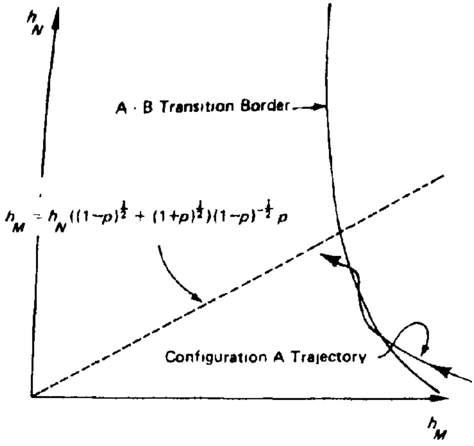
APPENDIX

In this appendix we discuss two issues: first, the possibility of multiple equilibria and, second, the nature of equilibrium paths from initial positions that are not on the h_M axis.

Whenever Nash equilibrium is the solution concept, as in this paper, the question of possible multiple equilibria arises. We avoided discussing multiplicity in the text by considering only the path with the maximum search and breach. If everyone else stops searching, the remaining individual must obviously find search unprofitable. Thus, taking any equilibrium path and altering it so that, at some arbitrary point, all individuals switch to Configuration D, we trivially generate a new equilibrium. (Of course, this change will require modification of earlier transitions.) Similarly, to the left of the line $h_M = h''_M$, an N's search is worthwhile only if other N's also search. Thus an equilibrium path following Configuration A between the line $h_M = h''_M$ and the A–C transition locus could, at any time, switch to Configuration C behavior and still remain an equilibrium trajectory. Indeed, an equilibrium path between these two curves could oscillate between Configurations A and C arbitrarily.

More interesting is the possibility of multiple equilibria involving A–B transitions. We have neither confirmed nor ruled out this possibility. If multiplicity were possible, a Configuration A trajectory would necessarily cross the A–B transition border as in Fig. A1. Anywhere on A–B transition border (more precisely, just to the left of the border), an N finds double breach unprofitable if everyone else follows Configuration B behavior. Therefore, any equilibrium path in Region A has an equilibrium continuation in Region B beginning at the A–B transition border. Suppose, however, that when an equilibrium path reaches the border, individuals persist with A behavior. The question is whether such behavior can be an equilibrium. The answer is yes if and only if the Configuration A trajectory from this point crosses the A–B transition border again.

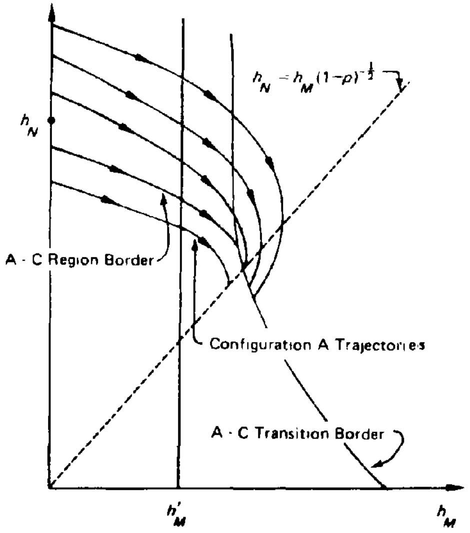
Now let us turn to equilibrium paths with initial positions not on the h_M axis. As long as the initial position lies below the line $h_M = h_N(1-p)^{1/2}$ (see (2)), the analysis is as before. Consider, therefore, the question of A–C tran-


h_N
A · B Transition Border
$h_M = h_N((1-p)^{\frac{1}{2}} + (1+p)^{\frac{1}{2}})(1-p)^{-\frac{1}{2}} p$
Configuration A Trajectory
h_M

A1

FIG. A1. Multiple equilibria at A–B transition border.

sitions when the initial position is above this line and when Region B does not exist. Fig. A2 shows the A–C transition border derived in the text (not yet shown, however, to be the region border in the present case) and a family of Configuration A trajectories. Moving backwards on one of these trajectories, the surplus from double breach remains positive as does an N's gains from search (see (16)). Therefore, Region A consists of all points to the left



A2

FIG. A2. A–C transitions for initial positions on the h_N-axis.

of the A–C transition border that lie above the trajectory just tangent to this border (see Fig. A2).

This analysis implies that positional values are not continuous in initial positions. As the initial position moves up the h_N axis, $V_N = X'$ until Region A is reached, at which point V_N increases discontinuously. The reason for this discontinuity is that equilibrium paths in Configuration A are impossible below Region A. Starting from a point just below this region, for example, an N's gain from search and double breach would be positive for awhile. However, if ever the configuration switches from A to C or D (as it must on an equilibrium path), the gains from search would be negative just before the transition, preventing such a path from being an equilibrium.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the NSF and mathematical assistance from Joel Yellin.

REFERENCES

DIAMOND, P. AND MASKIN, E., An equilibrium analysis of search and breach of contract. I. Steady states, *Bell J.* **10** (1979).

MORTENSEN, D., Specific capital, bargaining, and labor turnover, *Bell J.* **9** (1978), 572–586.

JOURNAL OF ECONOMIC THEORY 25, 196–218 (1981)

Equilibrium Distributions of Sales and Advertising Prices over Space

YVES BALCER

Department of Economics, University of Wisconson, Madison, Wisconson 53706

Received July 3, 1979; revised March 13, 1980

In "Equilibrium Distributions of Sales and Advertising Prices," Butters has established the existence of equilibrium price distributions for a market where all buyers and sellers are at the same location. The purpose of this article is to extend his model to incorporate location: sellers and buyers are spread uniformly on a surface, and sellers send advertisements (hereafter, ads) to potential buyers whose distance from the seller corresponds to a probability distribution. It will be shown that the equilibrium advertising prices do not always form a non-atomic distribution, and for very likely sets of parameters the distribution reduces to either one or two prices (exponential case) and to a finite number of prices (uniform case). However, the costs to the buyers, purchase price plus transportation cost, are always represented by a non-atomic distribution.

But why do the equilibrium distributions of advertisement prices and sales at sellers' prices reduce to a point? It is because an ad at the lowest possible price will not yield a sale with probability one, since an ad at price p received by a consumer at a certain distance from the seller is perceived as an ad at price p plus the transportation cost, and it is possible that an ad at $p' > p$ from a closer seller leads to a smaller total cost for the consumer, thus to a sale.

Though the model and all the results in this article are discussed in terms of location, they could also be interpreted in the context of markets for differentiated products, which could be manufactured at the same unit cost (for example, clothing of different colors, food of different tastes). Consumers are characterized by the specific property of the good they wish to purchase. Consumers buy one unit of the good for which they received an ad and whose combination of price and divergence from their preferred choice (measured in dollars) is the smallest. For this interpretation, the difference between Butters' and the present work is even more striking: When consumers are all alike, there is diversity in the behavior of producers (i.e., a continuum of prices are advertised) as in Butters; when consumers are

0022-0531/81/050196-23$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

different (as characterized by their location), the producers all behave identically (i.e., a unique price is advertised).

In the first section, we review the model and the results of Butters. In the second section, his model is extended to incorporate location. In the third section, an equilibrium is derived when the distance between a buyer reached by an ad and the seller follows an exponential distribution. In the fourth section, the case of uniform distribution for the distance between buyer and seller is examined. In the fifth section, it is shown that the competitive equilibrium in this market is inefficient with generally too few ads, as the reduction in transportation cost generated by the ads is not captured by the firm. Moreover, the socially efficient operation of this market can be achieved through corrective pricing, always in the uniform case, and sometimes in the exponential case. Finally, in the Appendix, a constructive proof of the general solution to the exponential case is provided; for a proof of the unicity of this solution, see [1].

1. BUTTERS' MODEL AND RESULTS

Model. The following assumptions hold:

(i) A single homogeneous good is being traded for money.

(ii) There are many identical buyers and sellers.

(iii) Sellers can send advertising messages informing buyers of their price and location. Sellers are not constrained to advertise the same price to all buyers.

(iv) Buyers have no other source of information about sellers. They receive ads. At the end, they must buy exactly one unit or forfeit their opportunity to buy.

(v) All buyers have the same limit price, m, namely, the maximum they are willing to pay for the good.

(vi) The cost to every seller of sending out each ad is b and each ad reaches a potential buyer. There are no economies of scale in advertising.

(vii) Ads are allocated randomly among buyers and independently of one another. Buyers receive ads freely and cannot affect their probability of receiving them.

(viii) All sellers have the same constant cost of production p_0.

(ix) All sellers know the buyers' limit price and the distribution of prices advertised by other sellers.

(x) Sellers maximize their expected profits given the behavior of the other sellers while buyers minimize the price they pay, subject to the above restraints.

For mathematical simplicity, Butters further assumes that there are an infinite number of sellers and buyers. The criterion for equilibrium is as follows: For a given advertised price distribution, the expected profit of each seller is zero, and no seller would expect higher profits at a different price and/or at a different level of advertisement. Let $\pi(p)$ be the probability that an ad at price p will result in a sale. Let $\beta(p)$ be the number of ads per (potential) buyer sent at a price no greater than p.

Using the analogy that buyers are empty urns and the number of ads per buyer, $\beta(p)$, is like the ratio of balls to urn, the probability, $\pi(p)$, of a successful ad, i.e., an ad resulting in a sale, is equal to the expected fraction of empty urns. Feller has shown that when the ratio of balls to urn remains constant while their number goes to infinity, the expected fraction of empty urns is the exponential of minus the ratio. So,

$$\pi(p) = \exp(-\beta(p)). \tag{1}$$

In equilibrium, there is no price for which the expected revenue can be greater than the expected cost, because of free entry. Thus, from the assumptions, we get directly

$$\pi(p) = b/(p - p_0). \tag{2}$$

Also, if one advertises a price k equal to the cost of production plus the cost of an ad, he generates a sale with certainty, since no one else can offer a lower price without taking a loss. Obviously no price greater than the reservation price m of buyers can be advertised. In addition, assume there is an interval of prices (p_1, p_2) belonging to $(p_0 + b, m)$, such that no price in that interval is advertised; it will therefore pay the seller who advertises at p_1 to increase his price to p_2, as this will lead to no loss of sale and increased revenue on each sale, a contradiction. So, all prices are advertised in the following interval:

$$m \geqslant p \geqslant p_0 + b = k. \tag{3}$$

Then, from Eqs. (1), (2), and (3), the density of ads $\gamma(p)$ and sales, $\delta(p)$, is obtained:

$$\begin{aligned} \gamma(p) &= 0 && \text{for} \quad p \leqslant p_0 + b, \\ &= 1/(p - p_0) && \text{for} \quad p_0 + b \leqslant p \leqslant m, \\ &= 0 && \text{for} \quad m \leqslant p, \end{aligned} \tag{4}$$

$$\begin{aligned} \delta(p) &= 0 && \text{for} \quad p \leqslant b_0 + b, \\ &= b/(p - p_0)^2 && \text{for} \quad p_0 + b \leqslant p \leqslant m, \\ &= 0 && \text{for} \quad m \leqslant p. \end{aligned} \tag{5}$$

Integrating (4), the total number of ads per person, $\beta(m)$, is found to be equal to $\ln((m - p_0)/b)$.

Finally, a measure of the welfare gain, W, resulting from the existence of a market is defined as the value of the items bought minus the sum of the production and advertising costs. Since, when β ads per buyer are already sent at a price lower than m, the probability of an additional successful ad is $\pi = \exp(-\beta)$; thus the fraction of buyers served is $(-\exp(-\beta))$. It then follows that

$$W = (1\text{-}\exp(-\beta))(m - p_0) - b\beta.$$

The optimal amount of advertising, β^*, is found by differentiating with respect to β and setting the expression equal to zero, yielding $\beta^* = \ln((m - p_0)/b) = \beta(m)$. The free market generates an optimal amount of advertising and the maximum possible welfare given the information technology.

2. EXTENSION OF BUTTERS' MODEL TO LOCATION

In order to incorporate locational considerations in Butters' model, the following assumptions must be added to the ten described in the previous section:

(ii′) Each point on an infinite surface is a location. Each location represents a different market, and the buyers and sellers are spread uniformly across the markets.

(vii′) The ads are allocated among each market according to a probability distribution, $F(t)$, which is a function of the distance t between the seller and the market where the buyer is located.[1] All ads are allocated independently.

(x′) The buyer purchases from the seller with the smallest total price, $p + ct$, where c is the unit transportation cost and t is the distance between the buyer and the seller.

The objective is to prove the existence and uniqueness of a symmetric Nash equilibrium and to derive the equilibrium distribution of prices paid and advertised. The structure of the argument will proceed as follows: assume first that an equilibrium exists, next derive the form it must take and check that it is truly an equilibrium. For mathematical simplicity, we shall assume that we have a continuum of markets, each with an infinite number of sellers and buyers. Our criteria for equilibrium are that at a given price distribution, the expected profit of each seller is zero, and that no seller would expect higher profits at a different price and/or at a different level of advertisement.

---

[1] Under the assumption of the model, the density of the location of a message sent from the origin i.e., the seller, is $f(x, y) = h(t)/2\pi t$, where $t = (x^2 + y^2)^{1/2}$.

By assumption (vii′), the density of ads at any point from a given source depends on the distance between that source and the point. Around each source, i.e., seller, there are circles of consumers who have an equiprobable chance of getting an ad; around each buyer, there are circles of sellers from whom he has an equiprobable chance of getting an ad. Integrating over those circles, one can solve the problem for a one-dimensional distribution. Therefore, without loss of generality, we shall limit our analysis to one-dimensional location problems, namely, the infinite line.

Since buyers pay the transportation cost, an ad at price p originating from a seller at a distance t from the buyer is considered by that buyer as an ad at price $p + ct$. So $\beta_S(p)$, the number of ads per potential buyer sent at price no greater than p, is different from $\beta_M(p)$, the number of ads per potential buyer received at a total cost (purchase price plus transportation cost) no greater than p. (The index M characterizes market prices [total costs] paid by the buyer, while the index S characterizes prices posted by the sellers.) Also, $\pi_M(p)$ is the probability that a buyer has no ad with a total cost smaller than p, while $\pi_S(p)$ is the probability that an ad at price p results in a sale. We immediately obtain the following relations, where $F(t)$ is the distribution of an ad at a distance t along the line[2]

$$\beta_M(p) = \int_0^{p/c} \beta_S(p - ct)\, dF(t), \tag{6}$$

$$\pi_S(p) = \int_0^{(m-p)/c} \pi_M(p + ct)\, dF(t). \tag{7}$$

The criterion for equilibrium, namely, expected revenue equals expected costs for a given price p, can now be written as

$$pq = p_0 q + bq/\pi_S(p), \tag{8}$$

where q is the quantity sold at price p. Also, by Butters' Propositions 2.1 (Eq. (2)), 2.2 (Eq. (3)), and 2.3 (Eq. (1)), we have the following relations:

$$\pi_M(p) = \exp(-\beta_M(p)), \tag{9}$$

$$\pi_S(p) \leqslant b/(p - p_0), \tag{10}$$

$$k \geqslant p_0 + b, \tag{11}$$

---

[2] Here again, the density of any point along the line from the origin of the message is given by $f(t) = \bar{f}(|t|)/2$. By symmetry restricting t to be positive, we double the integral and get Eqs. (6) and (7) as stated.

where relation (9) is true for all market prices, and relation (10) holds as an equality for all prices advertised by a seller. Finally, substituting Eqs. (6), (7), and (9) in (10), we get the crucial relationship between prices and the number of ads per potential buyer:

$$b/(p-p_0) \geqslant \int_0^{(m-p)/c} \exp\left(-\int_0^{(p/c)+c} \beta_S(p+c(t-u))\,dF(u)\right) dF(t). \quad (12)$$

Note that if $\beta_S(p)-\beta_S(p-\varepsilon)>0$ for arbitrarily small $\varepsilon>0$ at equilibrium, p is an advertised price and (12) is an equality at that point. Our present objective is the construction of a solution β_S which is non-decreasing in p for Eq. (12). Butters has solved this equation in two cases, when $F(0)=1$ (i.e., the markets are completely segregated, so the seller sends messages only to consumers at the same location as himself) and when the transportation cost is zero, i.e., the markets are completely integrated into a unique market without location. Without loss, p_0 is assumed to be zero, as this simply translates the reservation price and the supply price by p_0.

3. Exponential Distribution

In this section, it is assumed that the distribution of the distance between the origin of an add (i.e., the seller), and the recipient (i.e., the potential buyer) follows an exponential distribution of parameter λ. Without loss of generality, wee can assume that $\lambda=1$, since λ enters Eq. (12) always as c/λ.[1] Note that λ is a measure of the efficiency of the *ad* transmission to reach close-by buyers: an increase in transportation cost accompanied by a proportional increase in efficiency leads to an unchanged total cost for the potential buyer (closer buyers are reached but it costs them more per unit of distance to reach the store). Equation (12) is reduced to

$$b/p \geqslant \int_0^{(m-p)/c} \exp\left(-\int_0^{(p/c)+t} \beta_S(p+c(t-u))\exp(-u)\,du - t\right) dt. \quad (13)$$

Due to the difficulty of solving this equation for finite m, we first derive a solution to (13) when $m=\infty$. This solution is first obtained without the restriction that $\beta_S(p)$ must be non-decreasing. Again, the general solution will be adjusted to guarantee that this condition is met and finally extended to cover the case $m<\infty$.

---

[1] In Eq. (12), pursuant to the change of variables $v=\lambda u$ and $s=\lambda t$, we see that both c and λ always appear jointly as c/λ.

$m = \infty$, β_S *Unrestricted*

Assume that p is an advertised price in equilibrium. We can then replace π_S by its equilibrium value b/p. We obtain this solution to Eq. (7) by differentiation with respect to p,[4]

$$\pi_M(p) = \frac{b}{p} + \frac{bc}{p^2}, \tag{14}$$

provided $p \geqslant b(1 + (\sqrt{1 + 4c/d} - 1)/2) = k$. This condition is due to the fact that $\pi_M(p)$ is a probability, thus bounded by one. Another interpretation is; for p smaller than k to be an advertised p, given the small revenue it generates, the fraction of people without *ads* π_M in the closest markets (where this price p generates a total cost barely larger than itself) will have to exceed one, an impossibility. Note that when $c = 0$, the probability of making a sale in a given market $\pi_M(p)$ becomes equal to b/p, a result similar to Butters', reported in (2).

Using Eqs. (6) and (14), we obtain the following solution for (13) ignoring the restriction on $\beta_S(p)$, again by differentiation with respect to p,[5]

$$\beta_S(p) = \frac{2c}{p} - \frac{c}{p + c} - \ln\left(\frac{b}{p} + \frac{bc}{p^2}\right). \tag{15}$$

Note that when $c = 0$, $\beta_S(p) = -\ln(b/p)$ for $p \geqslant b$, which is Butters' result, obtainable by substituting (2) into (1). However, if c is relatively large, or b relatively small, $\beta_S(p)$ is not non-decreasing throughout the range. Thus it is not an admissible solution to our problem, because $\beta_S(p)$ decreasing is equivalent to saying that there are more *ads* at all prices smaller than p, than at all prices smaller than p', for some $p' > p$, an obvious impossibility. We shall come back to this point later.

For now, let us concentrate on the case when $\beta_S(p)$ is non-decreasing on $[k, \infty)$. A sufficient condition for that to be true is that $b/c \geqslant 2/3$.[6] Taking the derivative of (15) with respect to p, we obtain the density of *ads* at various (producer) prices,

---

[4] $b/p = (1/c) \int_p^{\infty} e^{p/c} e^{-u/c} \pi_M(u)\, du$. Taking the derivatives on both sides with respect to p, we obtain $-b/p^2 = -(1/c)\pi_M(p) + b/pc$.

[5] $-\ln(b/p + bc/p^2) = (1/c) \int_k^p e^{-p/c} e^{u/c} \beta_S(u)\, du$. Taking the derivatives on both sides with respect to p, we obtain $2/p - 1/(p + c) = (1/c)\beta_S(p) + (1/c)\ln(b/p + bc/p^2)$.

[6] The condition β_S is non-decreasing is equivalent to the condition γ_S, defined below in (16), is nonnegative. The sign of γ_S depends on the numerator given by (16). As a polynomial of the third degree, it has an inflexion point at $2c/3$ and a local minimum at $(-2 + \sqrt{10})c/3$. Since its value at $p = 0$ is negative, the polynomial takes negative values, then positive as p increases from 0 to infinity. In particular, we know that if $b \geqslant 2c/3$, then k is greater than 1.2c implying that $\gamma_S(p) > 0$ for all $p \geqslant k$. This follows from solving (14) for $\pi_M = 1$ and $p = 1.2c$.

$$\gamma_S(p) = \frac{p^3 + 2p^2c - 2pc^2 - 2c^3}{(p+c)^2p^2} \qquad \text{for} \quad p > k \tag{16}$$

with a mass of $2c/k - c/(k+c)$ at k. Note that when $c = 0$, the density of *ads* is simply $1/p$ for $p \geqslant b$, as established by Butters, see (4).

The density of sales at (seller) price p is the number of *ads* at price p times the probability that an ad will bring a sale:

$$\delta_S(p) = \gamma_S(p) \cdot \pi_S(p) = \gamma_S(p) \cdot \frac{b}{p} \qquad \text{for} \quad p > k, \tag{17}$$

with a mass of $(2c/k - c/(k+c))(b/k)$ at k. The density of sales at (buyer) price p is the number of ads at price p times the probability that an *ad* will bring a sale. By (9), we have[7,8]

$$\delta_M(p) = \frac{\partial \beta_M(p)}{\partial p} \cdot \pi_M(p) = \frac{b}{p^2} + \frac{2bc}{p^3} \qquad \text{for} \quad p \geqslant k. \tag{18}$$

Note again that when $c = 0$, $\delta_S(p) = \delta_M(p) = b/p^2$, as shown by Butters in (5). When p is very large, $\gamma_S(p)$ and both $\delta_S(p)$ and $\delta_M(p)$ approach Butters' results stated in (4) and (5), respectively.

The more difficult case, i.e., when not all prices greater than or equal to k are advertised in equilibrium,

$m = \infty$, β_S *Non-Decreasing*.

In that case, the formal solution developed in (15) is incorrect, but we shall use it to construct the correct answer to (13). We postulate the following solution, denoted $\beta'_S(p)$, to Eq. (13): $\beta'_S(p) = 0$ in $[0, k)$, $\beta'_S(p)$ is constant on $[k, \bar{k})$ and $\beta'_S(p) = \beta_S(p)$ on $[\bar{k}, \infty)$, where $\bar{k}$ is such that $\beta_S(p)$ is non-decreasing for $p \geqslant \bar{k}$. In other words, we postulate that only the following prices are advertised in equilibrium: k and all prices from $\bar{k}$ to ∞, a proof of which is given in the Appendix.

$m < \infty$.

In this model, the consumer will not purchase the good if the purchase cost plus the transportation cost exceed m, so $\pi_M(p) = 0$ for all $p > m$.

---

[7] Note that even though there is a concentration of *ads* at producer price k. However, since these *ads* reach at random a market, the dispersion is such that there is no concentration of *ads* with identical total costs to the consumers, i.e., the distribution of consumer prices has no mass.

[8] By (9), $\pi_M(p)\,\partial\beta_M(p)/\partial p = \pi_M(p)\,\partial(-\ln \pi_M(p))/\partial p = -\partial \pi_M(p)/\partial p$.

Moreover, since a seller does not know where an ad will land, he will not advertise at prices close to m.

Again, the idea for the solution is to construct it from the known solution of the infinite case and to truncate it at a given point, say $\bar{\bar{k}}$, beyond which no prices are advertised. This particular solution is postulated because the probability of a sale at market price $p < m$ is affected only by the *ads* issued at lower prices and because the probability of a sale at seller's price is directly affected only by the probability of a sale at market prices higher than p. Based on the first reason, the solution for the infinite case is still a solution for small p. From the second reason, we can assume that a reduction in advertisement at prices lower than m increases the probability of a sale per *ad* at all market prices to compensate for the finite reservation price. Truncation of the original solution is shown to be the right choice because the original and the new solution are identical for low p. This implies in particular that at the last advertised (seller) price $\bar{\bar{k}}$, the following relation holds, where the prime indicates the postulated solution for finite m:

$$\int_{\bar{\bar{k}}}^{\infty} e^{-u/c}\pi_M(u)\,du = \int_{\bar{\bar{k}}}^{m} e^{-u/c}\,\pi'_M(u)\,du, \tag{19}$$

by (7) and (10). This relation implies that $\pi_S(p) = \pi'_S(p)$ for $k \leqslant p \leqslant \bar{\bar{k}}$, and that $\pi_S(p) > \pi'_S(p)$ for $p > \bar{\bar{k}}$, since $\pi_M(p) < \pi'_M(p)$ in $(\bar{\bar{k}}, m)$. Thus the truncated solution $\beta'_S(p)$ associated with (20) solves (13) for finite m. Note that in the interval $(\bar{\bar{k}}, m)$

$$\beta'_M(p) = \beta_S(\bar{\bar{k}}) + \exp((\bar{\bar{k}} - p)/c)(\beta_M(\bar{\bar{k}}) - \beta_S(\bar{\bar{k}})).^{9} \tag{20}$$

Finally, the upper bound on the advertised prices, $\bar{\bar{k}}$, is independent of b when $\bar{\bar{k}}$ is larger than $\bar{k}$ and k.[10] Therefore, by (16), the density of *ads* is independent of their cost, with the exception of points k and $\bar{\bar{k}}$ where it has possibly two masses. The density of *ads* (4) derived by Butters is also independent of their cost, with the exception of the point where the density becomes positive.

Let $\tilde{m}$ be the smallest possible reservation price for which an equilibrium distribution of ads different from zero exists. It must satisfy the following relation,

$$b/k = \lim_{K\to 0}\{1 - \exp[K(\exp((k - \tilde{m})/c) - 1)]\}/K, \tag{21}$$

---

[9] By (6), $\beta'_M(p) = e^{-p/c}\int_0^{\bar{\bar{k}}} c^{-1}\beta_S(u)\,e^{u/c}\,du + e^{-p/c}\int_{\bar{\bar{k}}}^{p} c^{-1}\beta_S(u)\,e^{u/c}\,du$
$= e^{-p/c}\,e^{\bar{\bar{k}}/c}\,\beta_M(\bar{\bar{k}}) + e^{-p/c}\,\beta_S(\bar{\bar{k}})(e^{p/c} - e^{\bar{\bar{k}}/c})$.

[10] The left-hand side of (19) is equal to $b/\bar{\bar{k}}$ by (10). By (9), (14) and (15), Eq. (20) depends on b in the following fashion: $A + \ln b + e^{-p}B$. By (9), $\pi'_M(p)$ is equal b times of function of p. Therefore, b cancels from both sides of (19).

where the right-hand side is the integral of $\pi_M(p)$ defined by Footnote 17 from k to $\tilde{m}$. As $\tilde{m}$ gets smaller, the mass K at k is reduced to 0. Solving the above equation for $\tilde{m}$, we obtain

$$\tilde{m} = k + c\ln(k) - c\ln(k - b).$$

Recall that k is independent of $\tilde{m}$ by (14).

Summary

Three important price distributions of *ads* and sales were established in this section. The price distribution of *ads*, F_A, is characterized by 3 points, $k \leqslant \bar{k} \leqslant \bar{\bar{k}}$, 2 masses, K, $\bar{K}$, and a density function δ_S:

$$\begin{aligned} F_A(p) &= 0 && \text{if } p < k, \\ &= K && \text{if } k \leqslant p < \bar{k}, \\ &= K + \bar{K} && \text{if } p = \bar{k}, \\ &= \beta_S(p) = K + \bar{K} + \int_{\bar{k}}^{p} \delta_s(u)\,du && \text{if } \bar{k} \leqslant p \leqslant \bar{\bar{k}}. \end{aligned} \tag{22}$$

This distribution is such that under certain circumstances only one price is advertised ($k = \bar{k} = \bar{\bar{k}}$), under other circumstances only two prices are advertised ($k < \bar{k} = \bar{\bar{k}}$). These degeneracies happen when the reservation price is not exceedingly larger than the unit advertising cost. For relatively large reservation prices, there always exists a continuum of prices which are all advertised, between $\bar{k}$ and $\bar{\bar{k}}$.

The seller price distribution of sales, F_S, is

$$F_S(p) = \int_{k}^{p} b/u \, dF_A(u).$$

The market price distribution of sales, F_M, has no mass. In other words, there is no single price at which a large segment of transactions takes place:

$$\begin{aligned} F_M(p) &= 0 && \text{if } p < k, \\ &= 1 - \exp(K(\exp((k - p)/c) - 1)) && \text{if } k \leqslant p < \bar{k}, \\ &= F_M(\bar{k}) + \int_{\bar{k}}^{p} \delta_M(u)\,du && \text{if } \bar{k} \leqslant p \leqslant \bar{\bar{k}}, \\ &= F_M(\bar{\bar{k}}) + \exp(-\beta'_M(\bar{\bar{k}})) - \exp(-\beta'_M(p)) && \text{if } \bar{\bar{k}} \leqslant p \leqslant m. \end{aligned} \tag{24}$$

The main characteristics of these distributions have been computed numerically and are reported in Table 1. Since the distributions are

TABLE 1

Characteristics of the Price Distributions when the Production Cost $p_0 = 0$, the Exponential Distribution has Mean $\lambda^{-1} = 1$ and the Unit Transportation Cost $c = 1$

Parameters		Characteristics of price distribution				
Unit advertising b	Reservation cost m	Smallest advertised price k	Second advertised price $\bar{k}$	Largest advertised price $\bar{\bar{k}}$	Number of ads at price k K	Number of ads at price $\bar{k}$ $\bar{K}$
2.5	10	3.266	3.266	7.122	0.378	0.0
	6	3.266	3.266	3.602	0.378	0.0
	5	3.266	3.266	3.266	0.179	0.0
	4.716	3.266	3.266	3.266	0.0	0.0
1.0	10	1.618	1.618	7.122	0.854	0.0
	6	1.618	1.618	3.602	0.854	0.0
	3	1.618	1.618	1.618	0.530	0.0
	2.580	1.618	1.618	1.618	0.0	0.0
0.4	10	0.863	1.593	7.122	1.727	0.037
	6	0.863	1.593	3.602	1.727	0.037
	3	0.863	0.863	0.863	1.658	0.0
	1.486	0.863	0.863	0.863	0.0	0.0
0.1	10	0.370	4.184	7.122	3.599	0.206
	6	0.370	4.184	4.184	3.599	0.009
	3	0.370	0.379	0.370	3.475	0.0
	0.901	0.370	0.370	0.370	0.0	0.0

Note. All prices between $\bar{k}$ and $\bar{\bar{k}}$ are advertised.

independent of the choice of the numéraire, the results are reported only for $c = 1$. For example, if one is interested in the characteristics of the distribution when $c = 2.5$, $b = 1.0$ and $m = 10$, it suffices to compute the distribution for $c = 1$, $b = 0.4$ and $m = 4$ and multiply the resulting prices by 2.5. From Table 1, we see that the smaller the ratio of unit advertising cost, b, to transportation cost, c/λ, the more likely the distribution of advertised prices will reduce to one or two points for a given demand function characterized by the reservation price m.

4. UNIFORM DISTRIBUTION

In this section, it is assumed that the distribution of the distance between the origin of an ad and its recipient follows a uniform distribution with mean

$(2\lambda)^{-1}$.[11] Without loss of generality, we can again assume that $\lambda = 1$, since λ enters Eq. (12) always as c/λ. Equation (12) thus reduces to

$$\frac{b}{p} \geqslant \int_0^{\min(1,(m-p)/c)} \exp\left(-\int_0^{\min(p/c,\,1)} \beta_S(p + c(t-u))\,du\right) dt, \tag{25}$$

where β_S must be non-negative and non-decreasing. Also, for all advertised prices p (β_S strictly increasing at p), Eq. (25) is an equality.

Using a method similar to the one used in the previous section (see in particular Footnotes 4 and 5), it is easy to verify that no solution β_S to (25) may have pieces of continuity, as the continuity of β_S implies that β_S is decreasing. Contrary to the exponential case, it is very difficult to solve the problem for arbitrarily large reservation prices. Thus, we examine solutions that have a unique advertised price or finitely many for relatively small reservation prices.

Unique price. By postulating that the solution β_S has a unique point of increase (namely, a jump at k), the number of *ads* in the market at various prices is, by Eq. (4),

$$\begin{aligned} \beta_M(p) &= 0 && p < k, \\ &= \frac{K}{c}(p-k) && k < p < k + c, \\ &= K && k + c < p. \end{aligned} \tag{26}$$

Thus, at equilibrium, the following relation must hold between the advertised price and the number of *ads* sent per buyer by Eqs. (1) and (5).

$$b/k = (1 - \exp[-K(\min(m, k + c) - k)/c])/K. \tag{27}$$

However, this does not uniquely determine the solution. For β_S to be a solution, (25) must not be violated at prices different from k; a necessary conditon for this condition to hold is that the derivative of the right-hand side of Eq. (25) must be no greater than the derivative of its left-hand side:[12]

$$-b/k^2 = -(1 - \exp(-K))/c \qquad \text{if} \quad m > k + c \tag{28a}$$

$$= \text{in between those two values} \qquad \text{if} \quad m = k + c \tag{28b}$$

$$= -1/c \qquad \text{if} \quad m < k + c. \tag{28c}$$

---

[11] This particular value for the mean is used to simplify the equations in the rest of the section. However, to compare both distributions, they should have the same mean; this is accomplished by letting the unit transportation cost of the uniform case be twice that of the exponential case.

[12] When both (27) and (28.a) hold, the second derivative of the left-hand side must be no smaller than the one of the right-hand side at k. Otherwise, the derivative of the right-hand side of (25) will be greater than the one of the left for some $p > k$; consequently, (25) will be violated for some $p > k$. This relation holds exactly when b is relatively small.

TABLE 2

Advertised Price Distributions with Mass K at Price k

Parameters			Solution: (k, K)
$b/c < 1$	$m - c \leqslant \sqrt{bc}$	$m < 2\sqrt{bc}$	$(0, 0)$
		$m \geqslant 2\sqrt{bc}$	$(\sqrt{bc}, K')$
	$\sqrt{bc} < m - c < k^*$		$(m - c, K')$
	$k^* \leqslant m - c \leqslant \sim b\exp(K^*)^a$		(k^*, K^*)
	$\sim b\exp(K^*) \leqslant m - c$		More than one price
$b/c \geqslant 1$	$m - c \leqslant b$		$(0, 0)$
	$b \leqslant m - c \leqslant k^*$		$(m - c, K')$
	$k^* \leqslant m - c$		More than one price

[a] This relation holds exactly when b is relatively small.

Let (k^*, K^*) be the solution to (27) and (28a), and $(\sqrt{bc}, K')$ be the solution to (27) and (28c), where K' is determined by (27) for a given k. The solution is written for all combinations of parameters in Table 2.

Many prices. Though we cannot exhibit nor prove the existence of a multiprice solution in general, we can show that, with the exception of possibly one small price, there cannot be an interval more than c from one price to the next.[13]

If a solution exists, it must satisfy a set of conditions similar to (27) and (28a) that determine uniquely the solution. For example, the two price distributions, when $b/c \geqslant 1$ or when $b/c < 1$ and $b\exp(K^*) \leqslant m - c$, must satisfy the following equations:

$$k_2 = m - c,$$

$$\frac{b}{k_1} = \left(\frac{1}{K_1} - \frac{1}{K_1 + K_2}\right)\left(1 - \exp\left[-K_1\frac{(k_2 - k_1)}{c}\right]\right) + \frac{cb}{k_1^2(K_1 + K_2)},$$

$$\frac{b}{k_2}\exp\left[K_1\frac{(k_2 - k_1)}{c}\right] = \left(\frac{1}{K_1 + K_2}\right)\left(1 - \exp\left[-(K_1 + K_2)\frac{(k_1 + c - k_2)}{c}\right]\right) + \frac{1}{K_2}\left(1 - \exp\left[-K_1\frac{(k_1 + c - k_2)}{c} - K_2\right]\right),$$

$$\frac{cb}{k_1^2} = 1 - \exp\left[-K_1 - K_2\frac{(k_1 + c - k_2)}{c}\right].$$

---

[11] If the interval between two consecutive advertised prices is greater than c, the derivative of (25) at the first price k_1 will be of a form similar to Eq. (28a)(see Footnote 14). The second derivative of the left-hand side of (25) is smaller than the one of the right-hand side, when valued at k_1; this implies that for $p > k_1$, (25) is violated. This implies further that the last price is no less than $m - c$. Finally, an examination of the first derivative implies that the last price is no more than $m - c$; therefore, the last price is $m - c$.

Since the distribution of advertised prices does not exhibit very simple relationships with the unit cost transportation c, it is almost impossible to show that the solution to the location equilibrium converges with c to the non-location solution of Butters. However, it can be shown heuristically that this is the case in the following manner: For very small c, the solution to (25) is almost equal to a finite distribution of prices $k^*, k^* + c, \ldots, k^* + ic$ with masses $c/k^*, \ldots, c/(k^* + ic)$ at those points where $k^* = b + c/2$ and i is the integer part of $[(m - b)/c - 1]$.[14] The sum of those masses in the interval from b to $x \leqslant m$ converges as c goes to zero, to $\ln(x/b)$ the numbers of *ads* at prices less than or equal to x in Butters' model, completing the argument.

Summary

Three important price distributions of *ads* and sales were established in this section. The price distribution of *ads*, F_A, is a degenerate distribution with n points, such that no two consecutive advertised prices are more than c apart, with the possible exception of the first one (for simplicity only the one price distribution is described):

$$\begin{aligned} F_A(p) &= 0, & p &< k, \\ &= K, & k &\leqslant p. \end{aligned} \tag{29}$$

The seller price distribution of sales, F_S, is described in a similar way using (29) and (10):

$$\begin{aligned} F_S(p) &= 0, & p &< k, \\ &= Kb/k, & k &\leqslant p. \end{aligned} \tag{30}$$

The market price distribution of sales, F_M, has no mass and is obtained using (9) and (26):

$$\begin{aligned} F_M(p) &= 0, & & p < k, \\ &= 1 - \exp(-K(p - k)/c), & & k < p \leqslant \min(k + c, m), \\ &= 1 - \exp(-K), & & k + c < p \leqslant m. \end{aligned} \tag{31}$$

---

[14] Let (k_1, K_1) and (k_2, K_2) be the solutions to (27) and (28a) when the unit advertising cost is b and $b\exp(K_1)$, respectively. Substituting (28a) into (27), we obtain that $c = K_1 k_1 = K_2 k_2$. Replacing the value of K_1 and K_2, respectively, in (28a) and equating them, we obtain $k_1^2(\exp(c/k_1) - 1) = k_2^2(1 - \exp(-c/k_2))$. Expanding the exponential and simplifying, we get $k_1 + c = k_2 + c^2(k_2^{-1} - k_1^{-1})/6 - c^3(k_2^{-2} + k_1^{-2})/24 + o(c^3)$, where $o(c^3)$ is positive and little order of c^3. From the structure of the problem, the solution to (25) for (c, b, m) is simply $\{(ck_i, K_i)\}_{i=1}^{n}$ where $\{(k_i, K_i)\}_{i=1}^{n}$ is the solution to (25) for $(1, b/c, m/c)$. Thus, the solution for small c corresponds to a solution for large b. Since $k_2 \geqslant k_1 \geqslant b$, we can say that $k_1 + 1 = k_2$ and $ck_1 + c = ck_2$. If this held exactly, there would be mass points at ck_1, $ck_1 + c$, $ck_1 + 2c, \ldots$, with mass $1/k_1$, $1/k_1 + c$, $1/k_1 + 2c$ and the remark of Footnote 13 would not apply.

TABLE 3

Characteristics of the Price Distributions When the Ads are Sent According to a Uniform Distribution with Mean $\lambda^{-1}/2 = 1$, the Production Cost $P_0 = 0$ and the Unit Transportation Cost $c = 1$

Parameters		Characteristics of price distribution					
Unit advertising cost b	Reservation price m	First advertised price k_1	Second advertised price k_2	Third advertised price k_3	Number of ads at price k_1 K_1	Number of ads at price k_2 K_2	Number of ads at price k_3 K_3
2.5	8	solution has three weights					
	6	3.304	4.000		0.546	0.102	
	5	3.000			0.376		
	4.5	2.500			0.0		
1.0	10	solution has four weights					
	8	1.701	4.083	6.000	1.176	0.481	0.277
	6	1.701	4.000		1.176	0.4	
	3	1.414			0.295		
	2.828	1.414			0.0		
0.4	10	solution has four weights					
	8	0.955	4.088	6.000	2.094	0.480	0.276
	6	0.955	4.000		2.094	0.433	
	3	0.955			2.094		
	1.789	0.894			0.0		
0.1	10	0.450			4.446		
	6	0.450			4.446		
	3	0.450			4.446		
	0.894	0.447			0.0		

The main characteristics of these distributions have been computed numerically and are reported in Table 3. From Table 3 we see that for very low ratios of unit advertising cost, b, to expected transportation cost, c/λ, the price distribution reduces to only one point for relatively large reservation prices, m.

5. EFFICIENCY

In Butters, as we have shown in the first section, the competitive equilibrium is efficient, i.e., sellers send a socially optimal number of *ads*. The reason for this is that the marginal firm (the firm advertising the highest price) advertises only if the cost of an *ad*, b, is no greater than the expected revenue from it, $m\pi(m) = m \exp(-\beta(m))$. But, this is exactly the efficiency criterion: The cost of reaching an additional customer, $b/\exp(-\beta)$, given that β *ads* were sent already, should be no greater than the benefit of such a sale, m. However, in the presence of transportation cost, this result no longer holds. This is not surprising, as higher advertising levels result in two types of benefits: Greater volume of sale and lower transportation cost per sale, the latter not being captured by firms. We, therefore, expect the competitive equilibrium number of *ads* to be lower than is socially optimal.

Before jumping to the conclusion, we must consider another element. The efficient amount of advertising is determined by the benefit of an additional *ad*, while the competitive amount of advertising is determined by the average return of an *ad* (when all *ads* are sent at the same price): the usual effect in resource allocation under common property. This effect leads to higher than desirable levels of advertising, while the uncaptured savings on transportation costs by the producers push their decision in the opposite direction. The net impact, however, is that too few *ads* are sent: thus, in general our expectations will be correct (an exception is the uniform distribution case when $m = 2.0$, $b = 0.1$ and $c = 1.0$).

The efficiency problem is equivalent to the one of maximizing the benefits to consumers (value of the good reflected in the reservation price) minus advertising and transportation costs. Since all consumers are identical in all respects, there is no reason to charge different prices to different people, so a unique price is advertised. Moreover, once an *ad* has been sent, its cost is essentially a sunk cost; its social value lies in its ability to generate a sale of value m to the consumer for a transportation cost lower than m. Thus, the advertised price should be zero. The maximization problem can be written formally as follows:

$$\operatorname*{Max}_{K} \{mFR(K) - bK - cTR(K)\},$$

where K, FR, and TR represent respectively the number of *ads*, the fraction served, and the transportation cost. In the exponential case, the following holds:

$$FR(K) = 1 - \exp(-K(1 - \exp(-m/c))),$$

$$TR(K) = K \int_0^m u \exp(-u - K(1 - \exp(-u)))\, du.$$

by (24) when $k = \bar{k} = \bar{\bar{k}} = 0$.

For the uniform, the following is obtained from (31) when $k = 0$:

$$FR(K) = 1 - \exp(-K \min(m/c, 1)),$$

$$TR(K) = K \int_0^{\min(m/c,\, 1)} u \exp(-Ku)\, du.$$

For neither case is an explicit solution obtainable. Henceforth, the efficient quantity of *ads* and the resulting welfare level are computed numerically for some of the parameters used in Tables 1 and 3 and are reported in Tables 4 and 5 for the exponential and the uniform, respectively.

The remaining questions are, Can efficiency be reestablished in this market through corrective prices, (m', b')? If not, can, at least, welfare be improved by some corrective pricing? Before addressing these two issues, the problem is set as follows:

$$\operatorname*{Max}_{m', b'} \{mF_M(m'; m', b') - \int_0^{m'} u\,dF_M(u; m', b') - (b - b')\beta_S(m'; m', b')\},$$

where F_M, the integral, and β_S represent respectively the fraction served, the total cost (transportation and advertising), and the number of *ads* sent, as defined by (22) and (24) for the exponential and (29) and (31) for the uniform when the underlying parameters are m' and b'. Necessary conditions for the market to be efficient are (i) that a unique price, k', be advertised; and (ii) that the difference between the corrective reservation price, m', and the advertised price, k', be equal to the minimum of the reservation price, m, and of the maximum potential distance between the sender of an *ad* and its recipient. By Table 2 and (27), it is possible to choose $b/(m - c)$ such that the number of *ads* is efficient and the advertised price is $k = m - c$. Therefore, for the uniform, the first question is answered positively: Appropriate corrective pricing yields a competitive efficient solution. On the other hand, for the exponential it is generally not possible to satisfy both conditions simultaneously. For both distributions, corrective prices are reported in Tables 4 and 5. It is worth noting that both the reservation price, m, and the unit advertised cost, b, are taxed or subsidized in a non-intuitive fashion at the optimum. Our intuition is valid for small corrections around m and b: for both distributions, a small subsidy for the buyer per unit bought, $m' > m$, and/or a small subsidy for the seller on the unit price of an *ad*, $b' < b$, increase unequivocally the welfare resulting from the competitive market after accounting for the costs of the subsidies. Similarly, a small tax on unit transportation, $c' > c$, leads to a higher level of welfare. As in most tax problems, the direction of change in the prices around the original prices is not a good indication of the price distortions necessary to obtain an optimum level of welfare. A cursory examination of Tables 4 and 5, particularly the last two columns, should convince the skeptics.

TABLE 4

Welfare Levels under Different Market Conditions: Exponential Distribution

Parameters		Competitive market		Efficient		Corrective pricing[a]			
Unit advertising cost b	Reservation price m	Number of ads	Welfare level	Number of ads	Welfare level	Number of ads	Welfare level	Corrected ad cost b'	Corrected reservation price m'
2.5	10	1.073	3.333	1.375	3.518	1.367	3.445	1.138	6.641
	6	0.458	0.688	0.779	0.859	0.767	0.832	2.263	7.143
	5	0.179	0.155	0.540	0.376	0.534	0.364	3.309	8.067
1.0	10	1.990	6.008	2.390	6.226	2.445	6.190	0.218	4.142
	6	1.374	2.549	1.840	2.700	1.867	2.665	0.416	4.276
	3	0.531	0.291	0.939	0.431	0.937	0.428	1.105	4.470
0.4	10	2.906	7.720	3.469	7.933	3.487	7.930	0.108	5.280
	6	2.291	3.936	2.979	4.091	3.426	4.064	0.112	5.610
	3	1.658	1.254	2.219	1.316	2.421	1.310	0.223	4.131
0.1	10	4.292	9.067	5.329	9.180	5.380	9.180	0.042	10.000
	6	3.608	5.120	4.959	5.204	5.010	5.204	0.049	6.025
	3	3.566	2.204	4.470	2.232	4.478	2.232	0.061	3.050

[a] The last four entries in the columns (*ads*, b', m') are not very accurate as the function is very flat around the optimum.

TABLE 5

Welfare Levels under Different Market Conditions. Uniform Distribution

Parameters		Competitive Market		Efficient		Corrective Pricing			
Unit advertising cost b	Reservation price m	Number of ads	Welfare level	Number of ads	Welfare level	Number of ads	Welfare level	Corrected ad cost b'	Corrected reservation price m'
2.5	10	—	—	1.351	3.454	1.351	3.454	0.810	3.500
	6	0.648	0.812	0.753	0.829	0.753	0.829	1.867	4.719
	5	0.376	0.333	0.518	0.357	0.518	0.357	2.384	5.056
1	10	—	—	2.361	6.117	2.361	6.117	0.323	3.000
	6	1.612	2.531	1.791	2.611	1.791	2.611	0.519	3.500
	3	0.295	0.172	0.859	0.375	0.859	0.375	1.562	4.500
0.4	10	—	—	3.445	7.805	3.445	7.805	0.162	3.000
	6	2.527	3.879	2.922	3.968	2.922	3.968	0.221	3.000
	3	2.094	1.202	2.114	1.202	2.114	1.202	0.394	3.000
0.1	10	4.446	9.017	5.447	9.055	5.447	9.055	0.066	3.000
	6	4.446	5.064	5.069	5.076	5.069	5.076	0.077	3.000
	3	4.446	2.099	4.583	2.100	4.583	2.100	0.093	3.000

6. CONCLUSIONS

It was shown in this paper that for neither the exponential distribution nor the uniform one is the consumer surplus systematically larger (see Tables and 5). In the cases at hand, the exponential distribution has a larger standard deviation than the uniform distribution with the same mean, and also the exponential cuts the uniform only once and from above, implying a greater spread for the exponential. Though it is hard to generalize from only the two special cases analyzed here for the distance between a seller and a potential buyer, one can safely conclude that an increase in uncertainty does not necessarily lower consumer surplus, nor does it necessarily reduce the fraction of potential buyers actually served. Moreover, changes in basic parameters that lead to larger share of buyers served do not necessarily lead to larger consumer surplus. The only result corresponding to our intuition is that increases in basic costs (unit transportation or unit advertising) lower consumer surplus. For a discussion of these issues, see Balcer [1].

The distribution of advertised prices may be degenerate, depending on the relative values of unit transportation cost, unit advertising cost, and reservation price net of unit production cost. If one believes that in the real world unit transportation cost times the expected distance between potential buyer and seller, c/λ, is high relative to unit advertising cost, b, then only one or few prices are advertised, which seems more in line with casual observation.

Finally, the model implies that a competitive market will foster too little advertising when advertising is informative. Moreover, in many situations, optimal corrective pricing of unit advertising cost, purchase price and unit transportation cost is not sufficient to obtain the socially efficient allocation.

APPENDIX

In this appendix, we prove that the solution to the problem is non-decreasing in p and, in addition, is unique. To establish the first point, we postulate the following: $\beta'_S(p) = 0$ in $(0, k)$; $\beta'_S(p)$ is constant on $(k, \bar{k})$; and $\beta'_S(p) = \beta_S(p)$ on $(\bar{k}, \infty)$, where $\bar{k}$ is such that $\beta_S(p)$ is non-decreasing for $p \geqslant \bar{k}$.

The hypothesis that $\beta'_S(p)$ is equal to $\beta_S(p)$ for $p \geqslant \bar{k}$ implies that

$$\int_k^{\bar{k}} K\, e^{u/c}\, du = \int_k^{\bar{k}} \beta_S(u)\, e^{u/c}\, du, \text{[15]} \tag{A1}$$

---

[15] If this equation does not hold, $\beta'_M(p) \neq \beta_M(p)$ for some $p \geqslant \bar{k}$ by (6); so $\pi'_M(p) \neq \pi_M(p)$ by (9), thus $\pi'_S(p) \neq \pi_S(p)$ by (7), which implies that p could not be an advertised price for the modified solution, a contradiction to (10).

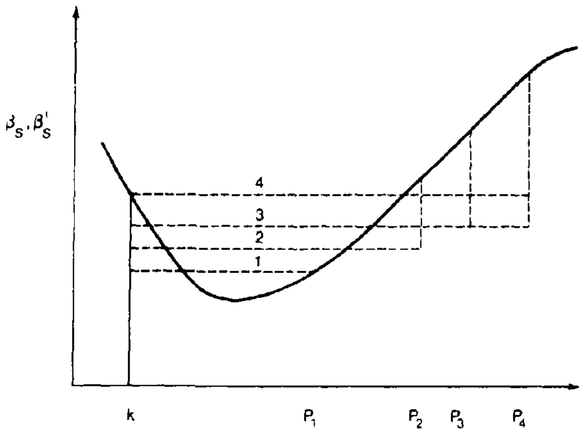


FIG. 1. Potential solutions for the number of ads at various seller's prices.

where $K = \beta'_S(k)$ is the number of *ads* at price k. Note that K is smaller than $\beta_S(\bar{k})$ and $\beta_S(k)$[16] and is increasing in $\bar{k}$ by (A1). By (A1), (6), (9) and (7), we have $\beta_M = \beta'_M$, $\beta_M = \beta'_M$ and $\beta_S = \beta'_S$ for all $p \geqslant \bar{k}$.

We shall present the rest of the argument graphically, but it can be formalized, since the exponential is a Polya distribution of class infinity. In

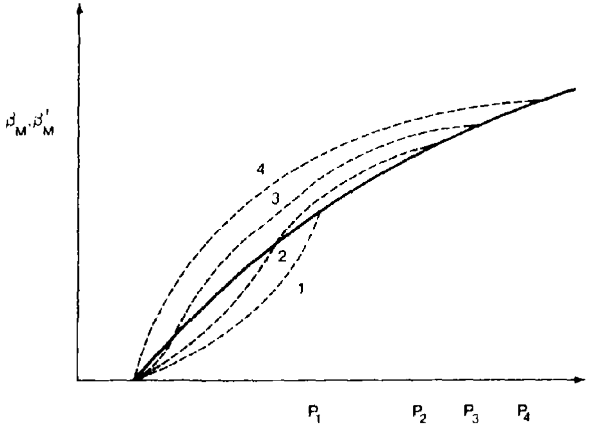


FIG. 2. Potential solutions for the number of a ads at various prices faced by the buyer, i.e., inclusive of the transportation cost.

---

[16] K is smaller than $\beta_S(\bar{k}) = \beta'_S(\bar{k})$ because β'_S is non-decreasing. If K was larger than $\beta_S(k)$, then $\beta'_M > \beta_M$ in the interval $(k, \bar{k})$ by (6) and (14); this implies by (9) that $\pi'_M < \pi_M$ in the same interval. Therefore, $\pi'_S(k) > \pi_S(k) = b/k$ by (7) and the fact that $\pi_M = \pi'_M$ for $p \geqslant \bar{k}$: a clear contradiction to (10).

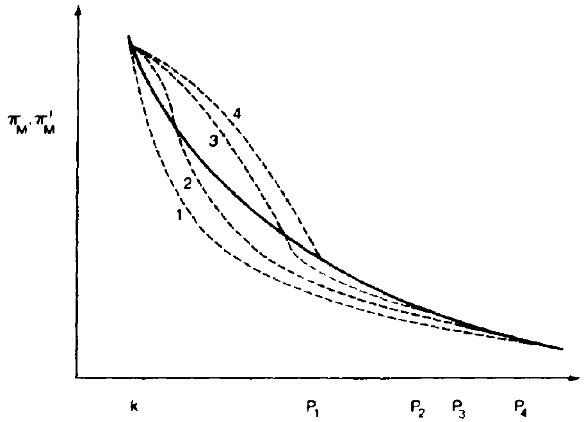


FIG. 3. Potential solutions for the probability of a sale at various prices faced by the buyer, i.e., inclusive of the transportation cost.

the four figures, the formal solution of (13), β_S, is illustrated by the solid line, and so are the associated functions β_M, π_M and π_S. As mentioned earlier, this solution does not meet the requirement that the number of *ads* is non-decreasing in p. Four possible candidates for solution are represented by the dashed lines: They all satisfy Eq. (A1), and the first is associated with the smallest K while the fourth is associated with the largest, as established in the previous paragraph. By Eqs. (6) and (A1) the functions $\beta'_M(p)$ are non-decreasing in K for all $p \geqslant k$, as depicted in the second figure. The third

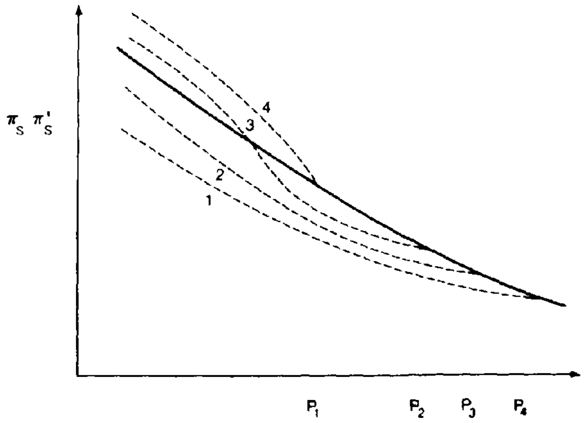


FIG. 4. Potential solutions for the probability of a sale at various seller's prices.

figure follows directly from (9). Finally, by (7) and since $\pi'_M(p)$ are non-increasing in K for all $p \geqslant k$, the same holds true for $\pi'_S(p)$; in addition, $\pi'_S(k) > \pi_S(k)$ for the smallest K feasible, and $\pi'_S(k) < \pi_S(k)$ for the largest K. Thus, there exists $\bar{k}$ such that $\pi'_S(k) = \pi_S(k)$ with $\pi'_S(p) = \pi_S(p)$ for $p \geqslant \bar{k}$. But is $\pi'_S(p)$ smaller than $\pi_S(p)$ in $(k, \bar{k})$? Yes; if not, $\pi'_M(p)$ would equal $\pi_M(p)$ more than once in $(k, \bar{k})$, which would contradict the fact that $\beta'_S(p)$ is equal to $\beta_S(p)$ no more than twice in $(k, \bar{k})$. To summarize, there exists $\bar{k}$ such that no *ad* takes place at prices in between k and $\bar{k}$ at equilibrium, and $\pi'_S(p)$ associated with the solution satisfies (10) at k and all p in $[\bar{k}, \infty)$.

In conclusion, the density of *ads* at various (seller) prices is $\gamma_S(p)$ for $p > \bar{k}$, as defined by Eq. (16), with a mass K smaller than $2c/k - c/(k + c)$ at k and a mass of $\bar{K}$ of $2c/\bar{k} - c/(\bar{k} + c) - \ln(b/\bar{k} + bc/\bar{k}^2) - K$ at $\bar{k}$. Note, there is no closed form formula for $\bar{k}$, it has to be computed numerically. We proceed along the same lines as before to derive $\delta_S(p)$. Finally, $\delta_M(p)$ is given by (18) for $p \geqslant \bar{k}$ and is equal to $-\partial\pi_M(p)/\partial p$ for $k \leqslant p < \bar{k}$, where $\pi_M(p) = \exp(K(\exp((k - p)/c) - 1))$.[17]

ACKNOWLEDGMENTS

This paper is an extension of the author's dissertation at the Massachussets Institute of Technology. He is particularly grateful to his advisor Peter Diamond, James A. Mirrlees and participants in the seminars at Queen's University, University of California, Berkeley and University of Wisconsin, Madison. Financial help from the Conseil National de Recherches du Canada is gratefully acknowledged.

REFERENCES

1. Y. BALCER, "Equilibrium Distributions of Sales and Advertising Prices over Space," Working Paper 7906, Social Systems Research Institute, University of Wisconsin, Madison, 1979.
2. G. Butters, equilibrium distribution of sales and advertising prices, *Rev. Econ. Studies* (1977), 465–491.
3. W. FELLER, "An Introduction to Probability Theory and Its Applications," Vol. I, Wiley, New York, 1950.

---

[17] $\pi_M(p) = \exp(-\beta_M(p))$ by (9) and $\beta_M(p) = K \int_0^{(p-k)/c} e^{-t}\, dt$ by (6).

JOURNAL OF ECONOMIC THEORY 25, 219–236 (1981)

Utility over Time: The Homothetic Case

TROUT RADER

Department of Economics, Washington University, St. Louis, Missouri

Received September 14, 1977; revised July 28, 1980

1. INTRODUCTION

The purpose of this paper is to show how the economics of choice over time can be strengthened by introduction of restrictive assumptions, namely, additivity and homotheticity of utility. We consider applications to labor supply, savings behavior, and long-run distribution and welfare measurement. Labor supply is mainly a function of the ratio of non-labor income to the wage rate. Savings generally increase with increases in the rate of interest. In the long run, distribution is not much dependent upon initial wealth, and instead is mainly determined by attitudes towards time. In the long run, the "as-if" social utility function is often well measured by consumer surplus. The applications appear to be strong. Therefore, it is appropriate to begin with a general argument for the main assumptions to be imposed.

Two early contributions to the economics of time are by Becker [1] and Koopmans [11]. The former presents a general framework without any specific assumptions about utility over time. At most, there is a homogeneous technology to produce consumption from goods and time (Pollak and Wachter [13]). Koopmans [11] adds highly intuitive assumptions to show utility additive and discounted over time. Koopmans' analysis could be extended to cases where utility is additive merely between remote time periods (see Rader [15, pp. 161–185]). This would correspond to the realistic case where the influence of distant consumption on preferences is negligible. Additivity must be reinterpreted in terms of a shifting environment of current consumption plus consumption sufficiently contiguous in time. The analysis following could be adapted to this most compelling case.

Merely adding utility over time does have implications for demand, for example, if complementaries occur between two goods they must occur between many others (e.g., [15, pp. 228–233]). However, the results are meager and insufficient for the kind of analysis economists generally want to apply. Therefore, we are led to add the assumption of homotheticity, namely,

0022-0531/81/050219-18$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

that increases in wealth leave relative consumption of goods unchanged. Clearly, this is a very special assumption. Nonetheless, in view of the failure to obtain strong results in general cases, the best strategy might be to begin with assumptions too strong and then weaken them without changing too much the basic theorems. Also, we know that homotheticity, however unrealistic, does allow for a variety of tastes. Therefore, some cases are covered that display generality beyond mere examples. As a further compensation, by adding homotheticity, the utility function can often vary over time so that explicit discounting is unnecessary.

2. BASIC CONCEPTS

We consider a preference relation $R \subset Y \times Y$ that ranks alternatives on Y. We write xRy for $(x, y) \in R$. Y is decomposable as a cartesian product,

$$\begin{aligned} Y &= X \times X \times \cdots \times X \\ &= \mathop{\times}_{i=1}^{n} X, \end{aligned}$$

where n is possibly infinite. R is *continuous* if for all $x \in Y$ the *contour* sets are closed,

$$\begin{aligned} R(x) &= \{y \mid yRx\}, \\ R^{-1}(x) &= \{y \mid xRy\}. \end{aligned}$$

Let $I \subset \bigcup_{i=1}^{n} \{i\}$ be a subset of integers less than or equal to n and denote $x_I = (x_i, i \in I)$, $x_{\sim I} = (x_i, i \notin I)$, $x = (x_I, x_{\sim I})$. Then R is *independent* if for all I, all x, y, preferences can be induced on x_I and y_I alone, independent of the value of the $\sim I$ coordinates: $(x_I, x_{\sim I})\, R\, (y_I, x_{\sim I})$ if and only if $(x_I, y_{\sim I})\, R\, (y_I, y_{\sim I})$. Let R_I denote R restricted to a given $x_{\sim I}$. R is *essentially independent* if also R_I is non-constant. If X is convex, then so is Y. R is (weakly) *convex* if, for all $x \in Y$, $R(x)$ is convex. If X is a cone, then so is Y. R is *homothetic if* Y is a cone and, for all $t > 0$, xRy if and only if $txRty$. R is *invariant* if for all (z, x), $(z, y) \in Y$, $(z, x)\, R\, (z, y)$ if and only if xRy. (Y must have infinite dimension for $R \subset Y \times Y$ to be invariant.)

A *utility representation* of R is a real valued function u such that

$$u(x) \geqslant u(y) \qquad \text{if and only if } xRy.$$

The function u is homothetic if the R so represented is homothetic.

3. UTILITY REPRESENTATION

We begin with the well-known additive utility result.

THEOREM 1 (Debreu–Leontief). *If R continuous,*[1] *essentially independent in three different I_i, $\times_{i=1}^{n} X^i$ is separable*[2] *and connected, then there exists a continuous u such that $\sum u_i(x_i)$ represents R. Furthermore, the u_i such that $\sum u_i$ represents R are unique up to an increasing affine transformation.*

Proof (Debreu [4]).

In analogy to the case where i refers to time, we term $u_i(x_i)$ (discounted) *temporal utility* and $\sum u_i(x_i)$ *intertemporal utility* or the *additive representation.* Our main result is as follows.

THEOREM 2. *If $\sum u_i(x_i)$ is monotonic, quasi-concave and homothetic in $(x_1, x_2, ...)$ and at least two u_i are non-constant, then there are concave linear homogeneous functions v_i and $1 \geqslant \beta \neq 0$, such that for all i, up to an additive constant, either*

$$u_i(x_i) = c_i v_i(x_i)^{\beta} \qquad (Bergson) \tag{1}$$

or

$$u_i(x_i) = c_i \ln v_i(x_i) \qquad (Bernoulli),\ c_i \geqslant 0. \tag{1'}$$

Remark 1. If $0 > \beta$, then $v_i(x_i)$ must be negative if u_i is to be monotonic. Furthermore, zero x_i gives negative infinite utility.

Let $U(x) \equiv \sum u_i(x_i)$, $x \equiv (x_1, x_2, ...)$.

COROLLARY 1. *$U(x)$ and $u_i(x_i)$ are either homogeneous of degree β (Bergson) or log homogeneous (Bernoulli).*

COROLLARY 2. *$U(x)$ is concave in x.*

Proof. The v_i's are concave as homogenous, quasi-concave functions and the u_i's are concave as concave functions of the v_i's (see Friedman [5]). Also, the sum of concave functions is concave. Q.E.D.

The theorem is notable since it specifies utility. The idea of its proof is to apply Bergson's [2] theorem to the case where scale multiples are made in consumption in a given period but the scale factor need not be identical over different periods. There results a specification that is Bergson or Bernoulli in scale effects.[3] That specification is then extended to the general case.

---

[1] If $n = \infty$, continuity in pointwise convergence is required (Rader [15, p. 158]).

[2] A topological space is *separable* if it has countable dense subset.

[3] In production theory, after transformation to a linear homogenous form, the Bernoulli function is termed Cobb–Douglas and the Bergson function, C.E.S. The elasticity of substitutes can be shown to be $1/(1-\beta)$, which exceeds one if $1 \geqslant \beta > 0$ and is positive and less than one if $\beta < 0$.

Proof. 1. *Bergson's Theorem.* We first show Theorem 2 when the x_i are real numbers. This gives Bergson's theorem [2; 19; 15, p. 212; 9, p. 31], but without explicit differentiability assumptions. The fact that Bergson's theorem is so general was announced in Rader [16, pp. 212–213], but the proof was inadequate. Also the possibility of $n = \infty$ was not mentioned, but as will be seen, finiteness of n plays no role in the proof.

2. *u is once differentiable.* Let u be monotonic, additive, quasi-concave, and homothetic, where $n \geqslant 2$, u_i non-constant for each i (or equivalently there are at least two distinct i's such that u_i is non-constant). We show $Du(x) = \sum Du_i(x_i)$ exists, and of course, $(Du(x))_i = Du_i(x_i)$.

According to Yaari [21], all but possibly one of the u_i's are concave. Let i be such that u_i is both concave and non-constant. Then for uncountable infinity of points x_i, $Du_i(x_i) \neq 0$. Also let $x_{\sim i}$ denote the vector x with the ith component deleted, $x = (x_i, x_{\sim i})$. It is easy to show that

$$g_c(x_{\sim i}) = \inf\{x_i \mid u(x_i, x_{\sim i}) = c\}$$

is convex in $x_{\sim i}$ (one needs monotonicity and quasi-concavity here). It follows that the *marginal rates of substitution* (*mrs*), $Dg_c^i(x_{\sim i})$, exists for almost all $x_{\sim i}$.

Also, u_i is monotonic and hence for all but a countable number of x_i, u_i is differentiable. Letting $c = u(x_i, x_{\sim i})$ a standard computation shows

$$(*) \quad (Dg_c(x_{\sim i}))_j = -\frac{Du_j(x_j)}{Du_i(x_i)}$$

whenever $Du_i(x_i) \neq 0$ and one of $(Dg_c(x_{\sim i}))_j$ or $Du_j(x_j)$ is defined (so that of course both are defined).

Let

$$S_i = \{x_i \mid Du_i(x_i) > 0\},$$

$$T_j = \{x_j \mid Du_j(x_j) \geqslant 0\}.$$

S_i is uncountable by concavity of u_i and, by the fact that u_j is monotonic, T_j contains all but a countable number points. Next define an arbitrary strictly positive vector $\bar{x}$. For an uncountable number of t, $t\bar{x}_j \in S_i$ and for all but a countable number of t, $t\bar{x}_j \in T_j$. Therefore, for an uncountable number of $t > 0$, both $t\bar{x}_i \in S_i$ and $t\bar{x}_i \in T_j$. Choose one such t. By $(*)$, $0 \geqslant Dg_c^i(t\bar{x}_{\sim i})j$. By homotheticity, the derivative above is constant in $t > 0$ and therefore exists for all $t > 0$. Furthermore since g_c is convex and monotonic, its derivatives are non-zero except possibly on the boundary where some quantities are zero.

Since the left-hand side of $(*)$ is defined for all $t\bar{x}$ and since $Du_i(\bar{x}_i) \neq 0$ exists for some i, some x_i, $Du_j(x_j)$ exists for all $x_j > 0$, $j \neq i$. Clearly

$Du_j(x_i) \neq 0$. For other points where u_i is not yet known to be differentiable, we have the computation

$$Du_i(x_i) = -\frac{Du_j(x_j)}{(Dg_c(x_{\sim i}))_j} \neq 0.$$

3. *u is twice differentiable a.e.* By homotheticity $Dg_c(x_{\sim i}) = Dg_{u(tx_i, tx_{\sim i})}(tx_{\sim i})$ for all t. By condition $(*)$, this translates to $Du_j(x_j)/Du_i(x_i) = Du_j(tx_j)/Du_i(tx_i)$ since $Du_i(tx_i) \neq 0$, $Du_i(x_i) \neq 0$. This implies $Du_j(tx_j) = Du_i(tx_i)(Du_j(x_j)/Du_i(x_i))$. For all $i \neq k$, k unique, u_i is concave and twice differentiable, a.e. Set $x_j = 1$ and consider the difference

$$Du_j(t) - Du_j(\bar{t}) = \frac{Du_j(1)}{Du_i(x_i)}(Du_i(tx_i) - Du_i(\bar{t}x_i)).$$

Pick any $t > 0$ and choose some x_i such that for $ty_i = x_i$, $Du_i^2(ty_i)$ exists (applying concavity of u_i). Then

$$\frac{Du_j(t) - Du_j(\bar{t})}{t - \bar{t}}$$

$$= \frac{Du_j(1)}{Du_i(x_i)} x_i \frac{Du_i(tx_i) - Du_i(\bar{t}x_i)}{(t - \bar{t})\, x_i} = \frac{Du_j(1)}{Du_i(x_i)} x_i \frac{Du_i(tx_i) - Du_i(\bar{t}x_i)}{tx_i - \bar{t}x_i}$$

and taking limits, $\bar{t} \to t$, $\bar{t}x_i \to tx_i = y_i$,

$$D^2u_j(t) \equiv \lim \frac{Du_j(t) - Du_j(\bar{t})}{t - \bar{t}}$$

$$= \frac{Du_j(1)}{Du_i(x_i)} x_i D^2u_i(y_i).$$

Since t was arbitrary, this proves twice differentiability a.e. for all u_j, $j \neq i$. Substituting j for i, we prove that u_i is twice differentiable.

4. *Evaluation of u_i.* Next comes the standard proof of Bergson's theorem. For all i except k, u_i is concave and absolutely continuous. By homotheticity $Dg_c^i(tx_{\sim i})$ is unchanged by small changes in t. Substituting $t(x_i, x_j)$ for (x_i, x_j), differentiating $(*)$ in t we have

$$0 = \frac{Du_j(x_j)}{(Du_i(x_i))^2} D^2u_i(x_i)\, x_i$$

$$- \frac{D^2u_j(x_j)}{Du_i(x_i)} x_j$$

or

$$(**) \quad x_i \frac{D^2u_i(x_i)}{Du_i(x_i)} = x_j \frac{D^2u_j(x_j)}{Du_j(x_j)}.$$

Since the left-hand and right-hand sides of (∗∗) are functionally independent, we have for some c, for all i,

$$(***) \quad x_i \frac{D^2 u_i(x_i)}{Du_i(x_i)} = c.$$

As is well known, solution of (∗∗∗) for u_i absolutely continuous gives, up to an additive constant,

$$\begin{aligned} u_i(x_i) &= x_i^{c+1} b_i, \qquad & c \neq -1, \\ &= b_i \ln x_i, \qquad & c = -1. \end{aligned}$$

We have the theorem in case the x_i are real valued $i \neq k$. Application of homotheticity gives the same for k.

5. *Extension to vector x_i's.* We specify the representation $\sum u_i(x_i)$, $u_i(0) = 0$ for all i. Consider scalars k_i and the representations, $x = (k_1 \bar{x}_1, k_2 \bar{x}_2)$, and

$$V(k_1, k_2, ...) = \sum u_i(k_i \bar{x}_i) \text{ (by Theorem 1).} \tag{2}$$

By Bergson's theorem, up to an additive constant, either

$$V(k_1, k_2, ...) = \sum_i c_i k_i^\beta, \qquad 1 \geqslant \beta \neq 0 \text{ (Bergson)}, \tag{3}$$

or

$$V(k_1, k_2, ...) = \sum_i c_i \ln k_i \qquad \text{(Bernoulli)}. \tag{3'}$$

In the Bergson case, we equate the independent parts of (2) and (3)

$$u_i(k_i \bar{x}_i) = c_i k_i^\beta = u_i(\bar{x}_i)\, k_i^\beta \qquad \text{up to an additive constant.}$$

Ostensibly the β depends upon the initial $\bar{x}_i$. However, it must be the same for any other x_j, since otherwise $\sum u_i$ would not be homothetic. Since $\bar{x}_i$ and $\bar{x}_j$ are independent, the same β applies to differing x_i's and x_j's.

We conclude that u_i is homogeneous of degree β. Setting $v_i(x_i) \equiv u_i(x_i)^{1/\beta}$ we have v_i linear homogeneous, quasi-concave such that

$$v_i(x_i)^\beta = u_i(x_i).$$

It follows that v_i is concave as well as quasi-concave (see Friedman, [5]).

In the Bernoulli case, we equate the independent parts of (2) and (3′) to obtain

$$u_i(k_i x_i) = c_i \ln k_i \qquad \text{up to an additive constant.}$$

Also to preserve homotheticity u_t must be log homogeneous in any other $\bar{x}_t$. Setting $v_t(x) \equiv \exp u_t(x_t)/c_t$ it follows that v_t is linear homogeneous, quasiconcave and therefore concave. Also, $u_t(x_t) = c_t \ln v_t(x_t)$. Q.E.D.

For application in Section 6, we need to further specify utility. To this purpose, we will assume invariance.

THEOREM 3. *If, in addition to the hypothesis of Theorem* 1, *R is invariant, then there exists u and an* $\alpha > 0$ *such that, for discounted temporal utility* u_t, $u_t(x_t) \equiv \alpha^t u(x_t)$.

Proof. Koopmans [11] or Rader [15, pp. 162–163].

COROLLARY. *Under the hypothesis of both Theorems* 2 *and* 3, *either*

$$u(x_1, x_2, \ldots) = \sum_t \alpha^t v(x_t)^\beta, \qquad 1 \geqslant \beta \neq 0 \ (\textit{Bergson}). \tag{4}$$

or

$$u(x_t, x_2, \ldots) = \sum_t \alpha^t \ln v(x_t) \qquad (\textit{Bernoulli}). \tag{4'}$$

4. APPLICATION I: LABOR SUPPLY

We consider a consumer who chooses between time at work and time at play (consumption). Let the vector of goods consumed at play be denoted x_1, goods at work, by x_2 and suppose utility is additive over time. Assuming no impatience[4], preferences are represented by

$$U(x_1, x_2) = \underbrace{u_1(x_1) + \cdots + u_1(x_1)}_{\text{number of times at play}} + \underbrace{u_2(x_2) + \cdots + u_2(x_2)}_{\text{number of times at work}}$$

$$= (N - T)\, u_1(x_1) + T u_2(x_2),$$

where N is the total number of periods and T is the number of periods at work. Dividing by N, we have the utility function to be studied

$$\frac{U(x_1, x_2)}{N} = (1 - t)\, u_1(x_1) + t u_2(x_2). \tag{5}$$

where t is the percentage of time at work. The function in (5) represents preferences if preferences are convex, homothetic, independent over time, but without impatience, and invariant in play and work, respectively, but not generally invariant in both. We also assume that the time is continuously gradated so that by an appropriate limit of time periods, Eq. (5) can be applied to evaluate the effect of decisions including at least t and x_1.

---

[4] There is no impatience if (x, y) is indifferent to (y, x).

The consumer is limited by his *budget constraint*,

$$px_1 \equiv wt + W, \tag{6}$$

where w is the wage rate and W is non-wage income. The first-order condition for constrained maximization of utility are necessary and sufficient,

$$(1-t)\,Du_1(x_1) = \lambda p, \tag{7}$$

$$\varnothing \equiv u_1(x_1) - u_2(x_2) = \lambda w, \qquad \lambda > 0, \qquad \text{or} \qquad \varnothing/w = \lambda > 0. \tag{8}$$

Substituting (8′) into (7) we have

$$(1-t)\,Du_1(x_1) = p\varnothing/w$$

or, post multiplying by x_1,

$$\begin{aligned}(1-t)\,Du_1(x_1)\,x_1 &= \varnothing/w\,px_1,\\ &= \varnothing/w(wt + W)\\ &= \varnothing t + \varnothing W/w\end{aligned}$$

or

$$t(Du_1(x_1)\,x_1 + \varnothing) = -\varnothing W/w + Du_1(x_1)\,x_1. \tag{7′}$$

Solving for t, *we have the labor supply equation*

$$\begin{aligned}t &= \frac{-\varnothing W/w}{Du_1(x_1)\,x_1 + \varnothing} + \frac{Du_1(x_1)\,x_1}{Du_1(x_1)\,x_1 + \varnothing}\\ &= \frac{-W/w}{1 + Du_1(x_1)\,x_1/\varnothing} + \frac{1}{1 + \varnothing/Du_1(x_1)\,x_1}.\end{aligned}$$

($\lambda > 0$ implies $\varnothing > 0$ is used here.) Note that the coefficient for $-W/w$ and the constant sum to one. The interpretation is that $1/(1 + \varnothing/Du_1(x_1)_1)$ is the percentage of time at work at zero wealth. To find the total time at work we subtract the product of the remaining fraction and the wealth-wage ratio. It is notable that the major independent influences on labor supply, wage rates and non-wage income figure prominently in the empirical labor supply studies such as that of Kalachek and Raines [8]. (See also Heckman [6] and Knieser [10].)

Equation (9) has the appearance of linearity in W/w. However, the term $x_1 \varnothing / Du_1(x_1)$ appears in the coefficients and it is generally dependent on W and w. Further specificity is gained by assuming preferences to be convex and homothetic in consumption at work and at play. Following Theorem 2, we have $u_1(x_1)$ is homogeneous of degree β (Bergson) or log homogeneous (Bernoulli) and

$$\begin{aligned} Du_1(x_1)\,x_1 &= \beta u_1(x_1) \qquad &\text{(Bergson)} \\ &= c_l \qquad &\text{(Bernoulli).} \end{aligned}$$

A further gain is made by assuming that preferences are homothetic in consumption at play and time at play, i.e., homothetic jointly in x_1 and $(1-t)$. This is the very special case where increased wealth, ceteris paribus, leads to equiproportional increases in consumption and leisure time. Then $u_2(x_2) = 0$ and $\varnothing = u_1(x_1)$, so that

$$t = \frac{1}{1 + u_1(x_1)/Du_1(x_1)} - \frac{1}{1 + Du_1(x_1)\,x_1/u_1(x_1)}\,\frac{W}{w}. \tag{9$'$}$$

For most cases, utility is Bergson[5] and labor supply takes on the very simple form

$$t = \frac{1}{1 + \dfrac{1}{\beta}} - \frac{1}{1+\beta}\,\frac{W}{w}. \tag{9$''$}$$

Comparing Eqs. (9′) and (9″), we have a paradox; more generality in utility leads to more specificity in labor supply.

It is interesting to see that if $\beta = 1$, then $t = 1/2 - 1/2(W/w)$. We may interpret this as saying (among other things) that if wealth is zero, the individual devotes half his waking time to work. Subtracting 8 hrs for sleep, this gives the customary 8-hr day.

Later (Theorem 6) we show that as a country increases in wealth, the dominant consumer type is one for which $\beta = 1$, if any such exists. Consequently, dominant consumers in rich economies will keep the 8-hr day. For very poor ones, the dominant consumer type will be seen to have a minimal β and will work less than in the rich economy. For example, for β greater than but nearly equal to zero, work time goes to zero! (If $-1 < \beta < 0$, there will be a boundary solution at no work at all).

---

[5] The Bernoulli functions are limits of the Bergson functions as $\beta \to 0$, so that the Bernoulli functions are in effect only one case of an infinity.

5. APPLICATION II: SAVINGS AND CONSUMPTION

The general *problem of demand* is to maximize $\sum u_i(x_i)$ subject to $\sum p_i x_i = \Theta$, where p_i and x_i are vectors and Θ is *wealth*. The resulting $x_i = x_i(p, \Theta)$ is the *quantity demanded*. We now interpret i to refer to time i, x_i to be consumption in time i, and p_i to be either current prices for time i goods as determined on future markets or, alternatively, discounted prospective prices for time i goods.

The main topic of the theory of demand is analysis of the change in demand with a change in price. This is complex even with the simplification to homotheticity and, in any case, does not normally yield an explicit solution of quantity in terms of price. We shall be content with a demand subproblem that admits an explicit solution.

In study of savings, we are not much concerned with the general question of demand behavior. Instead, we focus on the particular changes in discounted prices implied by changes in interest rates. These changes change future discounted prices of different goods in the same proportion. Therefore, we are led to consider changes that leave relative prices fixed at each time period.

Prices are of the form

$$p_i = \phi_i \bar{p}_i,$$

where ϕ_i is a scalar and $\bar{p}_i$ a fixed vector. From Theorem 2, each u_i is homothetic so that quantities demanded vary in proportions as prices do,

$$x_i(d_i) = d_i \bar{x}_i \qquad [15, \text{ p. } 215, \text{ Problem C}].$$

Thus, the budget constraint is $\Theta = \sum p_i x_i = \sum \phi_i p_i d_i \bar{x}_i$. Setting $q_i = \phi_i \bar{p}_i \bar{x}_i$ (and $\phi_i = q_i / \bar{p}_i \bar{x}_i$), we have that the consumer maximizes $\sum u_i(d_i \bar{x}_i) = \sum d_i^{\beta} u_i(\bar{x}_i)$ subject to $\Theta = \sum q_i d_i$, a problem in "prices" q_i and "quantities" d_i. This is the well-known problem of demand under a Bergson or Bernoulli utility function. The solution is explicit:

THEOREM 4.

$$d_i = \Theta \left(\frac{c_i}{q_i}\right)^{1/(1-\beta)} \left[\sum_j c_j^{1/(\beta-1)} q_j^{\beta/(\beta-1)}\right]^{-1}, \qquad (10)$$

where $q_i \equiv u_i(\bar{x}_i)$.

Proof. The Bernoulli case, $\beta = 0$, is easy and quite familiar.

For the Bergson case, the analysis is also standard. However, the computations are not so straightforward. We start from the first-order conditions for demand,

$$\beta c_i d_i^{\beta-1} = \lambda q_i. \tag{11}$$

We evaluate

$$\lambda = \beta \frac{U}{\Theta}. \tag{11$'$}$$

so that

$$c_i d_i^{\beta-1} = \frac{U}{\Theta} q_i. \tag{12}$$

Next, we compute (indirect) utility by summing over i,

$$\Theta = \sum q_i d_i = \sum (c_i)^{1/(1-\beta)} \left(\frac{U}{\Theta} q_i\right)^{1/(\beta-1)} q_i,$$

$$q_i = \left(\frac{U}{\Theta}\right)^{1/(\beta-1)} \sum c_i^{1/(1-\beta)} q_i^{1/(\beta-1)} q_i$$

or

$$\Theta^{\beta-1} = \frac{U}{\Theta} \left[\sum c^{1/(1-\beta)} q_i^{\beta/(\beta-1)}\right]^{\beta-1}$$

or

$$U = \Theta^{\beta} \left[\sum c_i^{1/(1-\beta)} q_i^{\beta/(\beta-1)}\right]^{1-\beta}. \tag{13}$$

Now substitute for U,

$$d_i^{\beta-1} = \frac{\Theta^{\beta} q_i}{\Theta c_i} \left[\sum (c_i)^{1/(\beta-1)} q_i^{\beta/(\beta-1)}\right]^{1-\beta}.$$

Applying the power $1/(\beta-1)$ gives the theorem. Q.E.D.

Wealth also depends on price and is computed as follows. Let the consumer produce goods y_i at time i with income $p_i y_i = \varphi_i \bar{p}_i y_i$ and define $\bar{p}_i y_i \equiv \bar{\Pi}_i$, $\bar{\Pi} = (\bar{\Pi}_1, \bar{\Pi}_2, \ldots)$. Then

$$\Theta = \sum p_i y_i = \sum \bar{\Pi}_i q_i$$

and

$$\frac{\partial d}{\partial q} = \left.\frac{\partial d}{\partial q}\right|_{\Theta} + \left.\frac{\partial d}{\partial \Theta}\right|_{q} \frac{\partial \Theta}{\partial q} = \left.\frac{\partial d}{\partial q}\right|_{\Theta} + \left.\frac{\partial d}{\partial \Theta}\right|_{q} \bar{\Pi}. \tag{14}$$

THEOREM 5. *If utility is Bernoulli or Bergson with $1 > \beta > 0$ and Θ increases with q_i, then an increase in q_i increases d_j for all $j \neq i$. If utility is Bergson with $0 > \beta$, then an increase in q_i may increase d_j if wealth increases fast enough, but otherwise d_j decreases, $j \neq i$.*[6]

COROLLARY. *An increase in the rate of interest decreases discounted wealth and discounted future prices so that savings increase if $1 > \beta > 0$ or if utility is Bernoulli.*

The Corollary verifies the classical view of savings whenever $\beta > 0$. In view of the critical role of a positive coefficient β, it is worth considering the likelihood of $0 > \beta$. To return to Remark 1 (following Theorem 2). Zero consumption in any period makes an individual infinitely ill off. This might be realistic over a long period, but, if we can apply our separability and homotheticity assumptions to a fine division of time, it is not realistic in period i alone.

6. APPLICATION III: ASYMPTOTIC SOCIAL WELFARE

Our third application is to many consumers, each with a utility function over time, $U^k(x^k)$, $k = 1,..., m$. *Social utility* or *welfare* is defined by

$$\Psi(x^1, x^2,..., x^m) = \sum_k b^k U^k(x^k), \tag{15}$$

where $b^k > 0$ is k's weight. We will mainly consider results independent of the particular value of the b^k's. Therefore, contrary to appearance, our analysis is not merely an exercise in the economics of cardinal utility. This point is deserving of emphasis and we devote an appendix to its exposition.

In anticipation of the Appendix, we assert that maximizing Ψ is equivalent to the distribution of goods in a way that is efficient in the Pareto sense. Hence, we define *efficient distribution* as the maximization of social utility $\Psi(x)$ subject to $\sum x^k = \bar{y}$, where $\bar{y}$ is vector of *available goods* and where the weights b^k are fixed but otherwise without significance. It should be noted that our use of efficient distribution is compatible with its use elsewhere, provided that we use sentences that make no mention of the weights. Then we can regard social utility as an analytic artifice without substantive content beyond the Pareto criterion.

The vector of consumption, x^k, is in the present context an infinite vector $x^k = (x_1^k, x_2^k,...)$. We also write $x^k = (x_t^k,\ t = 1, 2,...)$. Without further specification of the utility function, $U^k(x^k)$ is the sum of a long sequence of utilities over time. However, if k's preferences are invariant over time, then, by Theorem 3, $U^k(x^k)$ reduces to a relatively simple series,

---

[6] Theorem 5 is a solution to problem D in [15, p. 215].

$$U^k(x^k) = \sum [v^k(x_t^k)]^{\beta_k}(\alpha_k)^t. \tag{16}$$

The social utility is now of the form

$$\begin{aligned} \Psi(x) &= \sum_k b_k U^k(x^k) \\ &= \sum_k b_k \sum_t [v^k(x_t^k)]^{\beta_k}(\alpha_k)^t \\ &= \sum_t \sum_k b_k v^k(x_t^k)^{\beta_k}(\alpha_k)^t. \end{aligned} \tag{17}$$

If the α_k's were the same, then we could write

$$\begin{aligned} \Psi(x) &= \sum_t (\alpha_k)^t \sum v^k(x_t^k)^{\beta_k} b_k \\ &= \sum_t (\alpha_k)^t \psi(x_t^1,\dots,x_t^n), \end{aligned} \tag{17'}$$

where

$$\psi(x_t^1,\dots,x_t^n) = \sum v^k(x_t^k)^{\beta_k} b_k$$

is *temporal social utility* and $\Psi(x)$, *intertemporal social utility*. Finally, if all β_k's were equal, $\psi(x_t^1,\dots,x_t^n)$ would be homogeneous of degree $\beta = \beta_k$ and social utility homothetic. This would permit the use of consumer surplus (for consumer goods) over time to measure national welfare (see Rader [17], Katzner [9, p. 31], or Willig [20]). Equivalently, a Divisia index of consumption would measure welfare (Hulten [7], Rader [17]).

We derive conditions under which most consumption of a community of consumers adheres asymptotically to consumers with equal α_k's and equal β_k's. Explicitly, $\lim |x_t^k|/|x_t^j| = \infty$, whenever $(\alpha_j, \beta_j) \neq (\alpha_k, \beta_k)$ and (α_k, β_k) is the unique *dominant* type. Consequently, in the long run, social utility equals consumer surplus equals the Divisia index of consumption, provided only dominant types are given positive weight. It is as though long run allocation was accomplished with no consideration of non-dominant types. In effect, they are asymptotically invisible. (This does not imply that the social utility attributable to the non-dominant consumer disappears, since $\lim |x_t^k|/|x_t^j| = \infty$ or even $\lim |v^k(x_t^k)|/|v^j(x_t^j)| = \infty$ does not imply $\lim |v^k(x_t^k)|^{\beta_k}/|v^j(x_t^j)|^{\beta_j} = \infty$.)

Already, long-run behavior has been considered in Rader [14, Chap. 1; 16, Chap. 6]. Provided that consumption is bounded both from infinity and zero, all consumption ultimately goes to those having the highest discount factor α_k. In effect, the most patient consumers prevail, in keeping with the Biblical proverb "the meek shall inherit the earth."[7] However, meekness is not sufficient for our purposes. There may be several consumers with maximum α_k, each with different β_k's, whereupon only some k will dominate. In such cases, the temporal social utility ψ is not homothetic and consumer surplus equals Divisia index need not measure social welfare. There remains the possibility of more satisfactory measurement results whenever consumption is unbounded, either $y_t = \sum x_t^k \to 0$ or, for some good i, $(y_t)_i = \sum (x_t^k)_i \to \infty$.

THEOREM 6. *Suppose the hypothesis of Theorems 2 and 3, and also that there is efficient distribution and that in Eq.* (16), v_k *has a strictly positive derivative for all* $x^k \geqslant 0$, $x^k \neq 0$, *for all* k.[8]

(i) *Let* $\sum_k x_t^k \to 0$ *and suppose some consumer simultaneously has maximum* α_k *and minimum* β_k. *Then, for any other consumer* j *with* $(\alpha_j, \beta_j) \neq (\alpha_k, \beta_k)$,

$$\frac{|v^k(x_t^k)|}{|v^j(x_t^j)|} \to \infty$$

and

$$\frac{|x_t^k|}{|x_t^j|} \to \infty.$$

(ii) *Let* $\sum_k x_t^k$ *diverge and suppose some consumer simultaneously has maximum* α_k *and maximum* β_k. *Then, for any other consumer* j *with* $(\alpha_j, \beta_j) \neq (\alpha_k, \beta_k)$,

$$\frac{|v^k(x_t^k)|}{|v^j(x_t^j)|} \to \infty$$

and

$$\frac{|x_t^k|}{|x_t^j|} \to \infty.$$

---

[7] A very early result appears in Ramsey [18] where, given satiable utility, consumers, if any, with discount rate greater than the marginal product of capital attain bliss and all others, subsistence.

[8] For example, $V_k = \sum c_i x_i^\beta$, $c_i > 0$, $\beta > 0$. Otherwise, Bergson will not have strictly positive derivative so that the added condition is restrictive. Still, a large variety of other linear homogeneous functions are included.

By Theorem 6, the dominant consumers have $a_k = \max a_j$ and either β_k is a minimum or β_k is a maximum, depending upon whether the economy is poor or rich.

COROLLARY. *If all consumers are dominant over temporal utility* $\sum_k b_k v^l(x_t^l)$ *is homothetic and measured by consumer surplus.*

Proof. Rader [17].

Paradoxically, Theorem 6 shows that the ultimate position of consumers is relatively unaffected by their initial endowment of wealth. It may be objected that it may take a long time for the dominant consumers to predominate. However, a few examples should convince the reader otherwise. (In the example in Rader [14, Chap. 1], convergence to the dominant type is quick, indeed exponential.)

Part (i) shows that the predominant consumer type in poor economies is the most patient and the most stubborn. Suppose that zero utility at one time does not imply zero intertemporal utility, so that $\beta_k > 0$. Then, for k, such $\beta_k = 0$, i.e., Bernoulli utility over time, k would be the predominant consumer.

Part (ii) shows that dominant consumers in rich economies are those individuals with the most patience and the most flexible intertemporal preferences, i.e., highest β_k. Economically dominant citizens of rich economies are characterized by urbane flexibility. These would be the ones with $\beta_k = 1$, if any; i.e., utility linear over time.

The new hypotheses of Theorem 6 include the domain assumption that v^k has positive derivatives for all non-negative quantities. Among other things, this implies that temporal indifference curves are not asymptotic to the origin. In effect, no good is absolutely necessary for a non-zero level of temporal utility. The analytic role of the assumption is threefold. First, the domain assumption ensures that $Dv^k(x_t^k)$ is bounded from infinity. Since v^k is linear homogeneous, the derivative is zero homogeneous and a continuous zero homogeneous function on a closed set is bounded. Also, the domain assumption, together with the fact that $Dv^k(x_t^k)$ is strictly positive, implies that $Dv^k(x_t^k)$ is bounded from zero as well. Second, the domain assumption plus the strictly positive derivative ensures that $|v^k(x_t^k)|/|v^j(x_t^j)| \to \infty$ if and only if $|x_t^k|/|x_t^j| \to \infty$ and $|v^k(x_t^k)|/|v^j(x_t^j)| \to 0$ if and only if $|x_t^k|/|x_t^j| \to 0$. Third, infinite quantities of any good implies infinite temporal utility.

An example of $Dv^k(x_t^k) > 0$ for all $x_t^k > 0$ is where $v^k(x_t^k) = x_t^k$ is real valued or more generally $v^k(x_t^k) = \sum_i x_{t,i}^k c_i$, $c_i > 0$. Non linear examples do not include other Bergson or Bernoulli cases, as studied in the context of temporal utility. In effect, except for linear cases, we can disaggregate intertemporal utility no farther than the one period temporal utility function. This might be thought very restrictive. However, the set of linear homogeneous functions with strictly positive derivatives can be viewed as a convex set that is open and dense in the set of all linear homogeneous nondecreasing differentiable utilities. To see this, note that homogeneous functions defined on the nonnegative vectors are determined entirely by there behavior on the unit cube $\{x \mid x \geqslant 0,\ x_i \leqslant 1 \text{ for all } i\}$. The ones with strictly positive derivative are open in the C^1 topology on the unit cube. Furthermore any one with non-negative derivatives can be made to have strictly positive derivatives by adding an arbitrarily small linear function with positive derivatives.

Proof. The first-order condition for maximization of social utility implies, for $k \neq j$

$$b_k \alpha_k^t \, |v^k(x_t)|^{\beta_k - 1} \, Dv^k(x_t^k) = \lambda_t,$$

$$b_j \alpha_j^t \, |v^j(x_t)|^{\beta_j - 1} \, Dv^j(x_t^j) = \lambda_t.$$

Quantities bounded from infinity and zero include b_k, b_j, $Dv^k(x_t^k)$, $Dv^j(x_t^j)$. It follows that

$$\frac{\alpha_k^t \, |v^k(x_t^k)|^{\beta_k - 1}}{\alpha_j^t \, |v^j(x_t^j)|^{\beta_j - 1}}$$

is bounded from zero and infinity.

For $\alpha_k = \max_i \alpha_i$, $\alpha_j < \alpha_k$, we have

$$\frac{\alpha_k^t}{\alpha_j^t} = \left(\frac{\alpha_k}{\alpha_j}\right)^t \to \infty.$$

Therefore, for $\alpha_i > \alpha_k$

$$\frac{|v^k(x_t^k)|^{\beta_k - 1}}{|v^j(x_t^j)|^{\beta_j - 1}} \to 0$$

or

$$\frac{|v^k(x_j^k)|^{1 - \beta_k}}{|v^j(x_t^j)|^{1 - \beta_j}} \to \infty.$$

(i) Consider $\alpha_j = \alpha_k$, $\beta_j \neq \beta_k$. Let $\sum_i x_t^i \to 0$. Then $x_t^k \to 0$ and $|v^k(x_t^k)| \to 0$. If $\beta_j - \beta_k > 0$, then $|v^k(x_t^k)|^{\beta_j - \beta_k} \to 0$. Therefore

$$\frac{|v^k(x_t^k)|^{1 - \beta_j}}{|v^j(x_t^j)|^{1 - \beta_j}} \to \infty,$$

and we must have the conclusion of the theorem.

(ii) Let $\sum_t x_t^j$ diverge. Then for some j, $|v^j(x_t^j)| \to \infty$. If $\beta_j < \beta_k$ and $\alpha_k = \alpha_j$, then $1/|v^j(x_t^j)|^{\beta_k - \beta_j} \to 0$. Since

$$\frac{|v^k(x_t^k)|^{1-\beta_k}}{|v^j(x_t^j)|^{1-\beta_k}} \to \infty,$$

we have the conclusion of the theorem. Q.E.D.

APPENDIX: EQUIVALENCE OF SOCIAL WELFARE FUNCTIONS AND PARETO OPTIMALITY

Perhaps the most acceptable approach to social welfare is to consider the disaggregated utility vector $(V^1(x^1), \ldots, V^n(x^n))$. The utility vector is then evaluated by the weak inequality of Euclidean m-space $\geqslant$. This is the relation of *Pareto superiority*, or the Pareto welfare criterion.

The analysis will be relevant to the Pareto welfare criterion as much as to the more questionable social utility criterion. To see this, consider the *available goods*, $\sum x^k = \bar{y}$. The *Pareto criterion* leads to choice among those utility vectors that are maximal under the relation $\geqslant$ on the utility possibility set

$$S(\bar{y}) = \left\{ x \;\middle|\; \sum x^k = \bar{y} \right\}. \tag{16}$$

The social utility criterion is to maximize $\sum U^k(x^k)$ on $S(\bar{y})$. Provided that the U_k are concave and provided the particular affine transformation of U^k is unspecified, these are virtually the same.

Remark. If $b_k > 0$ for all k, the vector x maximizes $\Psi(x) \equiv \sum b_k V^k(x^k)$ on $S(\bar{y})$ only if the vector $(V^1(x^1), \ldots, V^n(x^n))$ is maximal on $S(\bar{y})$. Conversely, when the v^k are continuous, homothetic, and additive, and therefore concave, $(V^1(x^1), \ldots, V^n(x^n))$ is maximal on $S(\bar{y})$ only if there exist $b_k \geqslant 0$ such that $\Psi(x) = \sum b_k U^k(x^k)$ is maximized on $S(\bar{y})$.

Proof. A well-known implication of the linear hyperplane separation theorem for convex sets. Q.E.D.

Setting $U^k \equiv b_k V^k$, the remark shows how maximizing $\sum U^k$ is equivalent to maximal $(V^1, \ldots, V^m)$.

ACKNOWLEDGMENTS

The author has benefitted from meticulous comments of a referee, especially in regard to a proper proof of Theorem 2. Also, he gratefully acknowledges partial support of this research by the Department of Health, Education and Welfare, HEW/HCFA Contract 500-78-0055. Of course, the views expressed here are not necessarily those of HEW or HCFA.

REFERENCES

1. G. BECKER, A theory of the allocation of time, *Econ. J.* **75** (1965), 493–517.
2. A. BERGSON, Real income expenditure proportionality and Frisch's new methods of measuring marginal utility, *Rev. Econ. Studies* **4**, 33–52.
3. J. CHIPMAN AND J. MOORE, Why an increase in GNP need not imply an improvement in potential welfare, *Kyklos* **29** (1976), 391–418.
4. G. DEBREU, Topological methods in cardinal utility, *in* "Stanford Symposium on Mathematics in the Social Sciences" (Arrow *et al.*, Eds.), Stanford Univ. Press, Stanford, Cal., 1960
5. J. FRIEDMAN, Concavity of production function and non increasing returns to scale, *Econometrica* **41** (1973), 981–984.
6. J. HECKMAN, Shadow price, market wages, and labor supply, *Econometrica* **42** (1974), 679–694.
7. C. R. HULTEN, Divisia index numbers, *Econometrica* **41** (1973), 1017–1026.
8. E. KALACHEK AND F. RAINES, "Labor Supply of Low Income Workers," Technical Studies, President's Commission on Income Maintenance Programs, Washington, D. C., 1970.
9. D. KATZNER, "Static Demand Theory," McGraw–Hill, New York, 1971.
10. T. J. KNIESNER, An indirect test of complementarity in a family labor supply model, *Econometrica* **44** (1976), 651–670.
11. T. C. KOOPMANS, "Structure of Preference Over Time," Cowles Foundation Discussion Paper, No. 206 (1–66).
12. T. C. KOOPMANS, Representation of preference orderings over time, *in* "Decision and Organization" (McGuire and Radner, Eds.), North–Holland, Amsterdam, 1972.
13. R. POLLAK AND M. WACHTER, The relevance of the household production function and its implication for the allocation of time, *J. Pol. Econ.* **83** (1975), 255–278.
14. T. RADER, "Economics of Feudalism," Gordon and Breach, New York, 1971.
15. T. Rader, "Theory of Microeconomics," Academic Press, New York, 1972.
16. T. RADER, "Theory of General Economic Equilibrium," Academic Press, New York, 1972.
17. T. RADER, Equivalence of consumer surplus, the divisia index of output, and Eisenberg's addilog social utility, *J. Econ. Theory* **13** (1976), 58–66.
18. F. RAMSEY, A mathematical theory of savings, *Econ. J.* **38** (1928), 543-559.
19. P. SAMUELSON, Using full duality to show that simultaneously additive direct and indirect utilities implies unitary elasticity of demand, *Econometrica* **33** (1965), 781–796.
20. R. D. WILLIG, Consumer's surplus without apology, *Amer. Econ. Rev.* **66** (1976), 589–597.
21. M. YAARI, A note on separability and quasi-convexity, *Econometrica* **45** (1977), 1183–1186.

JOURNAL OF ECONOMIC THEORY 25, 237–254 (1981)

Direct and Indirect Weak Separability

W. HAQUE

University of Toronto, Toronto, Ontario, Canada

Received December 12, 1979; June 18, 1980

1. INTRODUCTION

In a well-known paper [6], Lau attempts to give necessary and sufficient conditions for both the direct and the associated indirect utility functions to be separable in quantity and price categories, respectively. His Theorem VIII says that direct and indirect additive separability implies that all category functions with the possible exception of a single category are homothetic; Theorem IX says that direct and indirect weak separability implies that all category functions are homothetic.

Additive separability is a special case of weak separability. Hence these two theorems contradict each other. As a matter of fact, Lau's Theorem IX is false. An immediate counterexample is the additive form (ii) of the direct utility function in his Theorem VIII; it yields an indirect utility function of the same form.[1]

With Theorem IX found false, one is naturally sceptical about the companion Theorem VIII. Lau stops the proof of this latter theorem half-way by declaring "The special case of (ii) may be proved in a way similar to that used in Theorem VII." Insightful as this remark is, a formal proof is necessary. Elsewhere [4], I verified Theorem VIII and strengthened it. As to Theorem IX, error creeps in at the very beginning of the proof. Faulty combinatorics starts with the assertion "Two possibilities are present" in [6, p. 390]. The untenable expression (iv-11) on p. 391 and using in the next expression price derivatives of the marginal utility of income as divisors, even when they may be zero as in the Samuelson counterexample, quickly leads to the false conclusion that all category functions are homothetic.

The purpose of the present paper is to establish the correct direct and indirect weak separability theorem. This requires the introduction of a mathematical concept—conditional homogeneity—and proving, for a special case, the converse of the "Chain Rule" of Calculus. Furthermore, as a by-

---

[1] This was earlier verified by Samuelson [8, Theorem, p. 356] for a special case of Lau's form (ii). As a referee informs me, this point was also made by Blackorby *et al.* [1, pp. 124–125].

0022-0531/81/050237-18$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

product, we shall show that direct and indirect weak separability while not admitting of Gorman's "perfect aggregators" does generate price indices with which the Strotz–Gorman budgetary problem [2, 10] can be solved.

2. PRELIMINARIES

Let $x = (x_1,..., x_n)$, $p = (p_1,..., p_n)$, I and $y = p/I = (y_1,..., y_n)$ represent consumption bundle, price vector, income and normalized price vector, respectively. $N = \{1,..., n\}$ is the index set and $\{N_1,..., N_T\}$ is a partition of N for $2 \leqslant T \leqslant n$. $\# N_s$ means the number of elements in N_s. x^s and y^s are subvectors of x & y having coordinates x_i & y_i, $i \in N_s$. $\bar{x}^s$ & $\bar{y}^s$ are the complements of x^s and y^s in x & y, respectively. $\Omega = \{z \in R^n \mid z_i > 0$, all $i \in N\}$ is both the commodity space and the normalized-price space. u is a direct utility function on Ω and $f(y) = (f^1(y),..., f^n(y))$ is the implied demand function. $v(y) \equiv -u(f(y))$ is the negative indirect utility function corresponding to u. $\lambda/I \equiv -\partial v(y)/\partial I$ is the marginal utility of income. η_i is the income elasticity of demand for good i. $I_s \equiv x^s y^s I$ is the expenditure on commodity group s. $\mathscr{G}$ is the graph of the demand function, i.e., $(x, y) \in \mathscr{G}$ iff $x = f(y)$, $y \in \Omega$. We make

ASSUMPTION 1. *u is strictly quasi-concave and has continuous first and second partial derivatives on Ω.*

ASSUMPTION 2. *The indifference hypersurfaces of u have nowhere-vanishing Gaussian curvature, i.e., the bordered Hessian matrix of u is nonsingular at all points of Ω.*

Strict quasi-concavity is known to imply monotonicity of u. Hence $u_i(x) > 0$, $i \in N$, $x \in \Omega$. Due to our assumptions the necessary and sufficient conditions for $\max_{x \in \Omega, yx \leqslant 1} u(x)$ are: $u_i(x) = \lambda y_i$, $i \in N$, $xy = 1$, yielding a unique solution $x = f(y)$, $\lambda = v(y)$, $y \in \Omega$, with the functions $f: \Omega \to \Omega$, $v: \Omega \to (0, \infty)$, continuously once differentiable. Furthermore, $v(y) \equiv -u(f(y))$, $y \in \Omega$, satisfies Assumptions 1–2 as u does. The first order conditions $v_i(y) = \lambda x_i$, $i \in N$ and $xy = 1$ uniquely solve the dual problem $\max_{y \in \Omega, xy \leqslant 1} v(y)$—the solution being $y = f^{-1}(x)$, $\lambda = \mu(x)$, $x \in \Omega$ with the functions $f^{-1}: \Omega \to \Omega$, $\mu: \Omega \to (0, \infty)$ having continuous first partial derivatives. The two maximization problems imply each other with the Lagrange multiplier

$$\lambda = \mu(x) = \sum_1^n u_i x_i = \sum_1^n v_i y_i \equiv v(y)$$

and $f: \Omega \to \Omega$ being a homeomorphism. For any subset X, Y of Ω, we have

$X = f(Y)$ iff $Y = f^{-1}(X)$. If X is open, then $f^{-1}(X)$ is open; if Y is open, then $f(Y)$ is open.

N_s is called weakly separable in u if the conditional ordering on the subspace of the vector x^s generated by u is independent of $\bar{x}^s$. Due to Assumption 1, this is equivalent to the independence of $u_i(x)/u_j(x)$ of $\bar{x}^s$, $i, j \in N_s$ (see Sono [9, Theorems 1–2]).

u is called weakly separable with respect to the partition $\{N_1, N_2, \ldots, N_T\}$ if N_s, $\forall s = 1, \ldots, T$, is weakly separable in u. This happens via Gorman [3, Lemma 1, pp. 387–388] if and only if

$$u(x) = F(u^1(x^1), \ldots, u^T(x^T)), \qquad x \in \Omega, \tag{1}$$

with F strictly monotone-increasing in each argument (for short, $F\uparrow$). Similarly, v is weakly separable with respect to the partition $\{N_1, \ldots, N_T\}$ if and only if

$$v(y) = G(v^1(y^1), \ldots, v^T(y^T)), G\uparrow, \qquad y \in \Omega. \tag{2}$$

Without loss of generality, we assume

$$\#N_s > 1, \qquad s = 1, \ldots, S,$$
$$\#N_s = 1, \qquad s = S+1, \ldots, T.$$

Hence, by normalization in (1), (2), we let

$$u^s(x^s) \equiv x^s, \qquad v^s(y^s) \equiv y^s, \qquad s = S+1, \ldots, T,$$

so that (1)–(2) can be re-written

$$u(x) = F(u^1(x^1), \ldots, u^S(x^S); x^0), F\uparrow, \qquad x \in \Omega, \tag{3}$$
$$v(y) = G(v^1(y^1), \ldots, v^S(y^S); y^0), G\uparrow, \qquad y \in \Omega. \tag{4}$$

where x^0, y^0 are the sub-vectors of x, y with coordinates x_i, y_i, $i \in \bigcup_{S+1}^T N_s \equiv N_0$. By relabelling,

$$N_s = \{s\}, \qquad s = S+1, \ldots, T.$$
$$N_0 = (S+1, \ldots, T),$$
$$x^0 = (x_{S+1}, \ldots, x_T),$$
$$y^0 = (y_{S+1}, \ldots, y_T).$$

The central question of this paper is: What are the necessary and sufficient conditions for both (3) and (4) to hold under Assumptions 1–2? We now prepare some background for an answer. First we point out some smoothness properties of (3)–(4) implied by Assumptions 1–2.

As argued in the just-cited Gorman Lemma, we can let

$$u^s(x^s) = u(\theta^1,\dots,\theta^{s-1},x^s,\theta^{s+1},\dots,\theta^S;\theta^0) \equiv u(x^s,\tilde{\theta}^s), \tag{5}$$

where $\theta = (\theta^1,\dots,\theta^S,\theta^0)$ is some reference vector in Ω. Hence, in view of (5) and Assumption 1, u^s is strictly quasi-concave, strictly monotone-increasing and twice-continuously differentiable. Due to the monotonicity of F and strict quasi-concavity of u, it follows that for fixed $x^s = a^s$, say,

$$F(u^1(x^1),\dots,u^{s-1}(x^{s-1}),u^s(a^s),u^{s+1}(x^{s+1}),\dots,u^S(x^S);x^0) \tag{6}$$

is strictly quasi-concave in $\tilde{x}^s$. By proving the converse of the Chain Rule of Calculus, under a special condition, we show, in Appendix B, that the partial derivatives $F_s(z)$, $F_{st}(z) \equiv F_{ts}(z)$, for $s, t = 1,\dots,T$ and $z = (u^1(x^1),\dots,u^S(x^S), x_{S+1},\dots,x_T)$ exist and are continuous in $x \in \Omega$. By analogy, all these smoothness properties of u^s and F also hold for v^s and G.

Let

$$z_s = u^s(x^s), \qquad w_s = v^s(y^s), \qquad s = 1,\dots,S,$$

$$z = (u^1(x^1),\dots,u^S(x^S);x^0), \qquad w = (v^1(y^1),\dots,v^S(y^S);y^0),$$

$$\mu^s(x^s) = \sum_{i\in N_s} u_i^s x_i, \qquad \nu^s(y^s) = \sum_{i\in N_s} v_i^s y_i,$$

$$\alpha^{ij}(x^s) = \frac{\mu_i^s(x^s)}{u_i^s(x^s)} - \frac{\mu_j^s(x^s)}{u_j^s(x^s)}, \qquad \beta^{ij}(y^s) = \frac{\nu_i^s(y^s)}{v_i^s(y^s)} - \frac{\nu_j^s(y^s)}{v_j^s(y^s)}, \tag{7}$$

where

$$\mu_i^s(x^s) = u_i^s + \sum_{l\in N_s} u_{il}^s x_l,$$

$$\nu_i^s(y^s) = v_i^s + \sum_{l\in N_s} v_{il}^s y_l.$$

If u is given by (3) and X is an arbitrary subset of Ω, then we will call u^s quasi-homothetic on X if for all $i, j \in N_s$, $u_i^s(x^s)/u_j^s(x^s)$ is homogeneous of degree zero on X or equivalently if $\alpha^{ij}(x^s) = 0$ on X for all $i, j \in N_s$. If X is open in R^n, then quasi-homotheticity will be synonymous with homotheticity and we can assume u^s to be homogeneous of degree 1 on X by appropriate normalization[2] in (3). Similarly, if v is given by (4), analogous statements hold verbatim for the category v^s on a subset Y of Ω. Let us now define[3] the vital concept of conditional homogeneity of degree zero.

---

[2] Since X is open, it is the union of a countable number of disjoint, open, connected sets X_l, $l = 1,\dots,\infty$. It is known that $\alpha^{ij}(x^s) = 0$, $\forall i$, $j \in N_s$, $x \in X_l$, iff $u^s(x^s) = \phi^l(\psi^s(x^s))$ with ϕ^l monotone-increasing and ψ^s homogeneous of degree 1. So (3) can be written as if u^s itself is homogeneous of degree one.

DEFINITION 1. A real-valued function F of class C^1 defined on an open subset Z of R^T is called "s-conditionally homogeneous of degree zero" (C_sH_0, for short) iff there exists a real-valued continuous function ξ^s of a real variable such that

$$\sum_{t\neq s} F_t(z)\, z_t = \xi^s(z_s)\, F_s(z), \qquad z \in Z.$$

F will be called C_sH_0I if it is C_sH_0 and strictly increasing in each argument.

For example, letting $z=(z_s, \tilde{z}_s)$, the function $[h(z_s) + \alpha \ln H(\tilde{z}_s)]^\beta$ (with h being a differentiable function of one variable and H linear homogeneous having first partial derivatives) is C_sH_0, where $\xi^s(z_s) = \alpha/h'(z_s))$. Notice that the usual definition of zero-degree homogeneity is a special case of the above with $\xi^s(z_s) = -z_s$.

3. INTERMEDIATE DISCUSSION

In this section, certain key results are assembled for answering our main question in the next section. Suppose, in (3), given $s \in \{1,\dots, S\}$, u^t, $t \neq s$, is homogeneous of degree 1 on some open subset X of Ω. Then

$$z_t \equiv u^t(x^t) = \sum_{i \in N_t} u_i^t x_i \equiv \mu^t(x^t) > 0, \qquad t \neq s,$$

$$\mu(x) \equiv \sum_1^n u_i x_i = F_s \mu^s(x^s) + \sum_{t \neq s} F_t z_t + \sum_{i \in N_0} F_i x_i,$$

so that F is C_sH_0 for $x \in X$ if and only if

$$\begin{aligned}\mu(x)/F_s - \mu^s(x^s) &\equiv \frac{1}{F_s}\left[\sum_{t\neq s} F_t z_t + \sum_{i \in N_0} F_i x_i\right] \\ &\equiv \xi^s(z_s), \text{ say,} \qquad \text{for} \quad x \in X, \end{aligned} \tag{8}$$

$\Leftrightarrow \mu(x)/F_s$ is independent of $\tilde{x}^s$ on the subset X.

---

[3] This definition is sufficient for the purpose of this paper. A referee has suggested that I give an equivalent definition in terms of a functional equation. This has inspired me to construct a complete theory of Conditional Homogeneity in Appendix C. I am greatly indebted to the referee for this invaluable suggestion.

Similarly, if in (4), for a given $s \in \{1,\dots, S\}$, v^t, $t \neq s$, is homogeneous of degree 1 on Y, then

$$w_t \equiv v^t(y^t) = \sum_{i\in N_t} v_i^t y_i \equiv v^t(y^t) > 0, \qquad t \neq s,$$

$$v(y) \equiv \sum_1^n v_i y_i = G_s v^s(y^s) + \sum_{t \neq s} G_t w_t + \sum_{i \in N_0} G_i y_i,$$

so that G is $C_s H_0$ for $y \in Y$ if and only if

$$\begin{aligned} v(y)/G_s - v^s(y^s) &\equiv \frac{1}{G_s}\left[\sum_{t \neq s} G_t w_t + \sum_{i \in N_0} G_i y_i\right] \\ &\equiv \eta^s(w_s), \text{ say}, \qquad \text{for} \quad y \in Y, \end{aligned} \tag{9}$$

$\Leftrightarrow v(y)/G_s$ is independent of $\bar{y}$s on the subset Y.

Notice that since F and G are monotone-increasing and $z_t > 0$, $w_t > 0$, $t \neq s$, (8)–(9) imply that $\xi^s(z_s) > 0$, $\eta^s(w_s) > 0$.

If u is given by (3), then the first order conditions for finding $\max_{yx \leqslant 1, x \in \Omega} u(x)$ are:

$$u_i(x) \equiv F_s(z)\, u_i^s(x^s) = \lambda y_i, \qquad i \in N_s, \qquad s = 1,\dots, s, \tag{10}$$

$$u_i(x) \equiv F_i(z) = \lambda y_i, \qquad i \in N_0, \tag{11}$$

$$yx = 1 \qquad [\text{or equivalently, } \lambda = \mu(x)]. \tag{12}$$

Similarly, if v is given by (4), the first order conditions for $\max_{xy \leqslant 1, y \in \Omega} v(y)$ are:

$$v_i(y) \equiv G_s(w)\, v_i^s(y^s) = \lambda x_i, \qquad i \in N_s, \qquad s = 1,\dots, S, \tag{13}$$

$$v_i(y) \equiv G_i(w) = \lambda x_i, \qquad i \in N_0, \tag{14}$$

$$xy = 1 \qquad [\text{or equivalently, } \lambda = v(y)]. \tag{15}$$

We now derive some results from these first order conditions.

3.1. *Direct Weak Separability*

Suppose u is given by (3) such that for some $s \in \{1,\dots, S\}$, there is an open subset X of Ω on which F is $C_s H_0$ and u^t, $t \neq s$, is homogeneous of degree 1. Fix an integer $k(t) \in N_t$, $t \neq s$. Then recalling (8), it is seen that (10)–(12), for $y \in f^{-1}(X)$, are equivalent to

$$\frac{u_i^s(x^s)}{\mu^s(x^s)+\xi^s(u^s(x^s))}=y_i, \qquad i\in N_s, \tag{16}$$

$$F_t u_j^t(x^t/x_{k(t)})=\lambda y_j, \qquad j\in N_t,\ t=1,\ldots,S,\ t\neq s, \tag{17}$$

$$\frac{u_i^t(x^t/x_{k(t)})}{u_j^t(x^t/x_{k(t)})}=\frac{y_i}{y_j}, \qquad i,j\in N_t, \tag{18}$$

$$F_i=\lambda y_i, \qquad i\in N_0, \tag{19}$$

$$\lambda=F_s\cdot(\mu^s(x^s)+\xi^s(u^s(x^s))). \tag{20}$$

We first show that (16) has a unique solution. Let $x^s=a^s$ be any solution of (16). Then Eqs. (17)–(20) with $x^s=a^s$ are the necessary and sufficient conditions for maximizing the strictly quasi-concave function (6) subject to the constraint $\sum_{i\notin N_s} y_i x_i \leqslant 1-y^s a^s$ and will therefore have a unique solution $\tilde{x}^s=\tilde{a}^s$, say. $a\equiv(a^s,\tilde{a}^s)$ is thus a solution of (10)–(12) and hence must be unique. It follows that (16) has a unique solution

$$x_i=f^{si}(y^s), \qquad i\in N_s, \tag{21}$$

or vectorially,

$$x^s=f^s(y^s), \qquad y\in f^{-1}(X). \tag{22}$$

Using (21) in the dual first order condition $v_i(y)=\lambda x_i$, we have

$$\frac{v_i(y)}{v_j(y)}=\frac{f^{si}(y^s)}{f^{sj}(y^s)}, \qquad i,j\in N_s, \qquad y\in f^{-1}(X). \tag{23}$$

Now (18) has the unique solution

$$\frac{x_i}{x_{k(t)}}=h^{ti}(y^t/y_{k(t)}), \qquad i\in N_t, \tag{24}$$

or, vectorially,

$$x^t/x_{k(t)}=h^t(y^t/y_{k(t)}) \tag{25}$$

since given $\tilde{x}^t$, (12) and (18) are necessary and sufficient conditions for uniquely maximizing the strictly quasi-concave function $u^t(x^t)$ subject to $y^t x^t \leqslant 1-\tilde{y}^t\tilde{x}^t$. Now using (24) in the condition $v_i(y)=\lambda x_i$, we have

$$\frac{v_i(y)}{v_j(y)}=\frac{h^{ti}(y^t/y_{k(t)})}{h^{tj}(y^t/y_{k(t)})}, \qquad i,j\in N_t, \quad t\neq s, \quad y\in f^{-1}(X). \tag{26}$$

Remark 1. If u is given by (3) subject to the restriction that there is an integer $t \in \{1,\dots, S\}$ and a subset X of Ω such that u^t is quasi-homothetic on X, then (10) implies that (24), and hence (26), holds, i.e., $v_i(y)/v_j(y)$, $y \in f^{-1}(X)$ is a zero-degree homogeneous function of y^t for all $i, j \in N_t$.

3.2. *Indirect Weak Separability*

Suppose v is given by (4) such that for some $s \in \{1,\dots, S\}$, there is an open subset Y of Ω on which F is $C_s H_0$ and u^t, $t \neq s$, is homogeneous of degree 1. Then arguing as above,

$$\begin{aligned} \frac{u_i(x)}{u_j(x)} &= \text{(i)} \quad g^{si}(x^s)/g^{sj}(x^s), \qquad i, j \in N_s, \quad x \in f(Y), \\ &= \text{(ii)} \quad \frac{g^{ti}(x^t/x_{k(t)})}{g^{tj}(x^t/x_{k(t)})}, \qquad i, j \in N_t, \quad t \neq s. \end{aligned} \tag{27}$$

Remark 2. If v is given by (3) subject to the restriction that there is an integer $t \in \{1,\dots, S\}$ and a subset Y of Ω such that v^t is quasi-homothetic on Y, then (27ii) holds, i.e., $u_i(x)/u_j(x)$, $x \in f(Y)$ is a zero-degree homogeneous function of x^t for all $i, j \in N_t$.

3.3. *Direct and Indirect Weak Separability*

Suppose u and v are given by (3)–(4). Logarithmically differentiating (10) and (13) with respect to y_k, $k \notin N_s$, $(x, y) \in \mathscr{G}$, we get

$$\frac{1}{u_i^s} \sum_{l \in N_s} u_{il}^s \frac{\partial x_l}{\partial y_k} = \frac{F_s}{\lambda} \frac{\partial}{\partial y_k}(\lambda/F_s), \qquad i \in N_s, \tag{28}$$

$$\frac{\partial x_i}{\partial y_k} = -x_i \frac{G_s}{\lambda} \frac{\partial}{\partial y_k}(\lambda/G_s), \qquad i \in N_s. \tag{29}$$

Using (29) in (28),

$$\left(\frac{1}{u_i^s} \sum_{l \in N_s} u_{il}^s x_l\right) \frac{\partial}{\partial y_k}(\lambda/G_s) = -\frac{F_s}{G_s} \frac{\partial}{\partial y_k}(\lambda/F_s), \qquad i \in N_s. \tag{30}$$

Similarly,

$$\left(\frac{1}{u_j^s} \sum_{l \in N_s} u_{jl}^s x_l\right) \frac{\partial}{\partial y_k}(\lambda/G_s) = -\frac{F_s}{G_s} \frac{\partial(\lambda/F_s)}{\partial y_k}, \qquad j \in N_s. \tag{31}$$

Subtracting (31) from (30) and using the notation in (7)

$$a^{ij}(x^s) \frac{\partial(\lambda/G_s)}{\partial y_k} = 0, \qquad i, j \in N_s, \quad k \notin N_s, \quad (x, y) \in \mathscr{G}. \tag{32}$$

Similarly, logarithmically differentiating (10) and (13) with respect to x_k, $k \notin N_s$,

$$\beta^{ij}(y^s)\frac{\partial(\lambda/F_s)}{\partial x_k}=0, \qquad i,j\in N_s, \quad k\notin N_s \quad (x,y)\in \mathscr{G}. \tag{33}$$

Remark 3. If u, v are given by (3), (4) and X is a subset of Ω, then, by Remarks 1–2, u^s is quasi-homothetic on X if and only if v^s is quasi-homothetic on $f^{-1}(X)$.

4. MAIN RESULT

The central question of this paper is now answered in the

DIRECT AND INDIRECT SEPARABILITY THEOREM. *The following are equivalent under Assumptions* 1–2.

(a) $u(x)=F(u^1(x^1),\ldots,u^S(x^S);x^0)$, $F\uparrow$, $\quad (x,y)\in\mathscr{G}$,

$v(y)=G(v^1(y^1),\ldots,v^S(y^S);y^0)$, $G\uparrow$.

(b) *There is a partition* $\{A_0,A_1,\ldots,A_S\}$ *of* Ω *such that* A_s, $s=1,\ldots,S$, *is an open subset of* R^n, A_0 *is closed relative to* Ω, $u(x)=F(u^1(x^1),\ldots,u^s(x^s));x^0)$, $F\uparrow$, $x\in\Omega$, *where* (i) u^s, $s=1,\ldots,S$, *is quasi-homothetic on* A_0, *and* (ii) F *is* C_sH_0, u^s *is not, but* u^t, $t\neq s$, *is, homogeneous of degree one, by normalization, on* A_s, $s=1,\ldots,S$.

(c) *There is a partition* $\{B_0,B_1,\ldots,B_S\}$ *of* Ω *such that* B_s, $s=1,\ldots,S$, *is an open subset of* R^n, B_0 *is closed relative to* Ω, $v(y)=G(v^1(y^1),\ldots,v^S(y^S);y^0)$, $G\uparrow$, $y\in\Omega$, *where* (i) v^s, $s=1,\ldots,S$, *is quasi-homothetic on* B_0 *and* (ii) G *is* C_sH_0, v^s *is not, but* v^t, $t\neq s$, *is, homogeneous of degree one, by normalization, on* B_s, $s=1,\ldots,S$.

Proof. Step 1. (a) ⇒ (b), (c).

Let

$$\begin{aligned}
A_s&=\{x\in\Omega \mid \alpha^{ij}(x^s)\neq 0,\ \text{some } i,j\in N_s\}, \qquad s=1,\ldots,S,\\
B_s&=f^{-1}(A_s), \qquad s=1,\ldots,S,\\
A_0&=\Omega-\bigcup_1^S A_s\equiv\bigcap_1^S(\Omega-A_s),\\
B_0&=f^{-1}(A_0)\equiv f^{-1}\left(\bigcap_1^S(\Omega-A_s)\right)\equiv\bigcap_1^S f^{-1}(\Omega-A_s)\\
&\equiv\bigcap_1^S(\Omega-B_s)\equiv\Omega-\bigcup_1^S B_s.
\end{aligned} \tag{34}$$

A_s, $s = 1,\dots, S$, is open in R^n because α^{ij} is continuous on the open subset Ω of R^n. On A_0, the continuous function α^{ij} takes the value zero and hence A_0 is closed relative to Ω. By the continuity of f, the set B_0 is closed relative to Ω and B_s, $s \neq 0$, is open in R^n.

We now demonstrate the properties of u, υ as stated in (b), (c).

First, on A_0, $\alpha^{ij}(x^s) = 0$, $\forall i, j \in N_s$, $s = 1,\dots, S$, so that (b)(i), (c)(i) follow from Remark 3.

On A_s, $s \neq 0$, by Remark 3, u^s, and hence υ^s, is not quasi-homothetic. This implies via (32)–(33) that for $s = 1,\dots, S$,

$$\frac{\partial[\nu(y)/G_s(w)]}{\partial y_k} \equiv \frac{G_s \nu_k(y) - \nu G_{st} \upsilon_k^t(y^t)}{G_s^2}$$

$$= 0, \qquad k \in N_t, \quad t \neq s, \quad y \in B_s. \tag{35}$$

$$\frac{\partial[\mu(x)/F_s(z)]}{\partial x_k} = 0, \qquad k \notin N_s, \quad x \in A_s. \tag{36}$$

From (35),

$$\frac{\nu_k(y)}{\upsilon_k(y)} = \frac{\nu G_{st}}{G_s G_t}, \qquad k \in N_t, \quad t \neq s, \quad y \in B_s. \tag{37}$$

Using (37) in the well-known equation[4]

$$\eta_k = 1 + \frac{\nu_k}{\upsilon_k} + \frac{1}{\nu} \sum_1^n \nu_l y_l \tag{38}$$

we see that all commodities in the set N_t have the same income elasticity. Hence, for all $i, j \in N_t$, $t \neq s$, $(x, y) \in A_s \times B_s$, $s = 1,\dots, S$, we have

$$0 = \frac{\partial}{\partial I} \frac{x_i}{x_j} \equiv \frac{\partial}{\partial I} \frac{\upsilon_i^t(y^t)}{\upsilon_j^t(y^t)} = -\frac{1}{I} \sum_{l \in N_t} y_l \frac{\partial}{\partial y_l} \frac{\upsilon_i^t(y^t)}{\upsilon_j^t(y^t)}, \tag{39}$$

which implies the quasi-homotheticity of υ^t, $t \neq s$, on B_s and hence of u^t, $t \neq s$, on A_s. This implies (since A_s, B_s are open) that u^t, υ^t, $t \neq s$, are homogeneous of degree 1 on A_s, B_s, respectively, by normalization. This together with (35)–(36) implies via (8)–(9) that F, G are $C_s H_0$ for $(x, y) \in A_s \times B_s$. We have shown that (b)(ii), (c)(ii) hold.

Finally, we show that $\{A_0, A_1,\dots, A_S\}$, $\{B_0, B_1,\dots, B_S\}$ are partitions of Ω. It is enough to show $A_s \cap A_t = \varnothing = B_s \cap B_t$, $s \neq t$, $s, t \in \{1,\dots, S\}$. Suppose $x \in A_s \cap A_t$. By (b)(ii) we get the contradiction: (a) $\alpha^{ij}(x^t) = 0$, $\forall i, j \in N_t$, since $x \in A_s$ and (b) $\alpha^{ij}(x^t) \neq 0$, some $i, j \in N_t$ since $x \in A_t$. This shows that

---

[4] Equation (38) is the equivalent form of Eq. (13) in Samuelson [7, p. 40].

$A_s \cap A_t = \varnothing$. If $y \in B_s \cap B_t$, $s \neq t$, then $f(y) \in A_s \cap A_t$ contradicting $A_s \cap A_t = \varnothing$. Hence $B_s \cap B_t = \varnothing$, $s \neq t$.

Step 2. (b) ⇒ (a), (c) ⇒ (a).

Given the partition $\{A_0, A_1, \ldots, A_S\}$, define $\{B_0, B_1, \ldots, B_S\}$ exactly by (34). The latter is a partition of Ω exactly by the last sentence of Step 1. Nowletting $X = A_0$, $t = 1, \ldots, S$ in Remark 1 and letting $X = A_s$, $s = 1, \ldots, S$, in (23) and (26), it follows that $v_i(y)/v_j(y)$ is independent of $\bar{y}^s$ for all $i, j \in N_s$, $s = 1, \ldots, S$, $y \in \Omega$. This means v is given by (4). We have shown (b) ⇒ (a).

Similarly, using (27) and Remark (2), (c) ⇒ (a).

Conclusion: By Steps 1–2, (b) ⇔ (a) ⇔ (c).

5. STROTZ–GORMAN BUDGETRY

As is well known, Strotz [10] and Gorman [2] assume a direct utility function given by our type (3). Their main point is to discover price indices $(P^1(p), \ldots, P^T(p))$ so that the sectoral income I_r, $r = 1, \ldots, T$, is representable as

$$I_r = J^r(P^1, \ldots, P^T, I), \qquad r = 1, \ldots, T. \tag{40}$$

In Gorman [2, Eq. (10), p. 471], it is found necessary and sufficient for "Local Aggregation," and hence necessary for "Perfect Aggregation" that $(\partial I_r/\partial p_i)/(\partial I_r/\partial p_j)$, $i, j \in N_t$, $r \neq t$, be independent of r. It is tacitly assumed that the partial derivatives $\partial I_r/\partial p_i$, $i \notin N_r$, do not vanish. This condition is not fulfilled by a preference ordering that is both directly and indirectly weakly separable. This is seenfrom the fact that our Eq. (22) gives $I_s = I y^s f^s(y^s)$ for $y \in B_s$, $s = 1, \ldots, S$, so that Gorman aggregation, local or perfect, is not possible in the strict sense. However, Gorman aggregation is possible in a slightly modified sense which simply replaces the single Gorman index $P^s(p)$ by two indices $v^s(y^s)$ & $\nu^s(y^s)$ on our sets B_s, $s = 1, \ldots, S$. This is seen as follows.

If u, v are given by (3)–(4), then by our Direct and Indirect Weak Separability Theorem, (9) holds for $y \in B_s$, $s = 1, \ldots, S$. Using (9) in (13)–(14), for $y \in B_s$, $s = 1, \ldots, S$,

$$\begin{aligned} &\text{(a)} \quad I_s = \frac{\nu^s(y^s)}{\nu^s(y^s) + \eta^s(v^s(y^s))} I, \\ &\text{(b)} \quad I_r = \frac{v^r(y^r)}{\nu^s(y^s) + \eta^s(v^s(y^s))} \frac{G_r(w)}{G_s(w)} I, \qquad r \neq s, \quad r = 1, \ldots, T, \end{aligned} \tag{41}$$

where $v^r(y^r) \equiv y_r$, $r \in N_0$.

Since $v^t(y^t) = v^t(p^t)/I$, $t \neq s$, (41b) has the form

$$I_r = J^r(v^s(y^s), v^s(y^s), (v^t(p^t))_{t \neq s}, I), \qquad r \neq s, \quad r = 1,\ldots, T, \tag{42}$$

where $v^t(p^t)$, $t \neq s$, is the so-called "Klein Aggregator," As to "Local Aggregation," define the "differential price indices"

$$\begin{aligned} \delta_t &= \sum_{i \in N_t} v_i^t(y^t)\, dp_i, \qquad t = 1,\ldots, T, \\ \sigma_s &= \sum_{i \in N_s} v_i^s(y^s)\, dp_i. \end{aligned} \tag{43}$$

Then taking the differential of I_s & I_r from (41) and using (43), we get the form, for $y \in B_s$,

$$\begin{aligned} dI_s &= \sigma_s a^s(p^s, I) + \delta_s b^s(p^s, I) + c^s(p^s, I)\, dI, \\ dI_r &= \sum_{t=1}^{T} \delta_t a^{rt}(p, I) + \delta_s a^{rs}(p, I) + b^{rs}(p, I)\, dI, \qquad r \neq s, \end{aligned} \tag{44}$$

which is Gorman's Eq. (6) in [2, p. 470] modified by an extra differential index σ_s for the sth commodity group.

For $y \in B_0$, Gorman aggregation—nay Klein aggregation—is, of course, possible due to the quasi-homotheticity of u^s, $s = 1,\ldots, T$, on the set A_0.

APPENDIX

A. LEMMA 1. *Suppose* X *is an open subset of* R^n; Y & Z *are subsets of* R^T & R^1. $\{N_1,\ldots, N_T\}$ *is a partition of* $\{1,\ldots, n\}$. *For* $x \in X$, $x = (x^1,\ldots, x^s,\ldots, x^T)$, *where* x^s *is the sub-vector consisting of elements* x_i, $i \in N_s$. *The functions* $f\colon X \to Y$, $g\colon Y \to Z$ *and* $h \equiv g \circ f\colon X \to Z$ *are so defined that* (a) $f(x) = (f^1(x^1),\ldots, f^s(x^s),\ldots, f^T(x^T))$ *for every* $x \in X$, (b) *for some* s, *there exists* $k \in N_s$ *such that* f^s *is continuous and strictly monotone in* x_k *on* X, *and* (c) $f_k^s(x^s)$ *exists at a given point* $x \in X$. *Then* $g_s(f(x))$ *exists if* $h_k(x)$ *does.*

Proof. Let $e_k = (0, 0,\ldots, 1, 0,\ldots, 0)$ be the kth unit vector in R^n. Then X being open, $\exists \eta > 0$ such that $x + \theta e_k \in X$ if $|\theta| < \eta$. With x given, the function $\psi(\theta) \equiv f^s(x^s + \theta e_k^s)$ is strictly monotone on the interval $(-\eta, \eta)$. Define the function

$$\begin{aligned} \phi(\theta) &= \frac{g(f(x + \theta e_k)) - g(f(x))}{\psi(\theta) - \psi(0)} \\ &= \frac{[h(x + \theta e_k) - h(x)]/\theta}{[f^s(x^s + \theta e_k^s) - f^s(x^s)]/\theta}, \qquad 0 < |\theta| < \eta. \end{aligned} \tag{45}$$

Suppose $h_k(x)$ exists. Then (45) implies

$$\lim_{\theta\to 0}\phi(\theta)=\frac{h_k(x)}{f^s_k(x^s)},$$

i.e., given $\varepsilon>0$, $\exists\delta\in(0,\eta)$, such that

$$0<|\theta|<\delta\Rightarrow\left|\frac{g(f(x+\theta e_k))-g(f(x))}{\psi(\theta)-\psi(0)}-\frac{h_k(x)}{f^s_k(x^s)}\right|<\varepsilon. \tag{46}$$

Let

$$\begin{aligned} z=\psi(\theta),\qquad \xi&=(f^1(x^1),\dots,f^{s-1}(x^{s-1}),z,f^{s+1}(x^{s+1}),\dots,f^T(x^T)),\\ \sigma_1&=\min\{\psi(-\delta),\psi(\delta)\},\\ \sigma_2&=\max\{\psi(-\delta),\psi(\delta)\}.\end{aligned}$$

Then $\psi(0)\equiv f^s(x^s)\in(\sigma_1,\sigma_2)$ and $\psi^{-}:(\sigma_1,\sigma_2)\to(-\delta,\delta)$ is a function due to the strict monotonicity and continuity of ψ on $(-\delta,\delta)$. Hence

$$z\neq f^s(x^s),\qquad z\in(\sigma_1,\sigma_2)\Leftrightarrow 0<|\psi^{-1}(z)|<\delta. \tag{47}$$

Equations (46)–(47) imply: given $\varepsilon>0$, $\exists\sigma_1$, σ_2 such that

$$z\neq f^s(x^s),\qquad z\in(\sigma_1,\sigma_2)\Rightarrow\left|\frac{g(\xi)-g(f(x))}{z-f^s(x^s)}-\frac{h_k(x)}{f^s_k(x^s)}\right|<\varepsilon,$$

i.e.,

$$g_s(f(x))=h_k(x)/f^s_k(x^s).$$

Remark 4. But for the extra assumption of monotonicity of f^s with respect to x_k, our hypothesis in Lemma 1 is the same as that of the Chain Rule of Calculus. Under our hypothesis, the Chain Rule says: if $g_s(f(x))$ exists, then $h_k(x)$ does. We have proved the converse at the cost, of course, of monotonicity of f^s in x_k.

To be sure, Lemma 1 generalizes and formalizes my 1974 version of the converse of Chain Rule (see [5, p. 453]). There I showed, for the transformation $V(y)=G(u^*(y))$, the existence of $G'(u^*(y))$ if $V_k(y)$ exists where u^* is the indirect utility function; in my proof, I only used the strict monotonicity of u^* in the kth argument. Recently, Blackorby, Primont and Russell [1, Theorem 3.5, pp. 78–79] gave another version under the additional, though unnecessary, assumption of directional differentiability of the utility function.

B. *Existence of F_s, F_{st}*

We have

$$u(x) = F(u^1(x^1),\dots, u^T(x^T)), u^t(x^t)) \equiv x_t, \qquad t = S + 1,\dots, T,$$

Since $u_i(x) > 0$, $u_i^s(x^s) > 0$, Lemma 1 yields

$$\begin{aligned} H^s(x) &\equiv F_s(u^1(x^1),\dots, u^t(x^t),\dots, u^T(x^T)) \\ &= u_i(x)/u_i^s(x^s), \qquad i \in N_s, \end{aligned} \tag{50}$$

whence

$$H_j^s(x) = \frac{u_i^s(x^s)\, u_{ij}(x) - u_i(z)\, u_{ij}^s(x^s)}{[u_i^s(x^s)]^2}, \qquad j \in N. \tag{51}$$

Equation (51) and the fact that $u_j^t(x^t) > 0$ applied to Lemma 1 guarantee the existence of F_{st}. Hence from (50)

$$H_j^s(x) = F_{st} u_j^t(x^t). \tag{52}$$

Equations (50)–(52) imply

$$\begin{aligned} F_{st} &= \frac{u_{ij}(x)}{u_i^s(x^s)\, u_j^t(x^t)} = F_{ts}, \qquad s \neq t, \\ F_{ss} &= \frac{u_i^s(x^s)\, u_{ij}(x) - u_i(x)\, u_{ij}^s(x^s)}{(u_i^s(x^s))^2\, u_j^s(x^s)}. \end{aligned} \tag{53}$$

Since $u_i(x)$, $u_{ij}(x)$ are continuous in x so are F_s, F_{st}, s, $t = 1,\dots, T$.

C. *Conditionally Homogeneous Functions*

In this section, we generalize the concept of homogeneous functions and the associated Euler Theorem and its converse.

Let M be a subset of $\{1,\dots, s,\dots, T\}$ and $M' = \{1,\dots, s,\dots, T\} - M$. Consider the system of ordinary differential equations

$$\frac{dz_s}{dt} = \frac{a^s(z_s)}{t}, \qquad t \neq 0, \qquad z = (z_1,\dots, z_s,\dots, z_T) \in Z, \tag{54}$$

where Z is an open set in R^T, the functions a^s, $s = 1,\dots, T$, are continuously differentiable and $a^s(z_s) \equiv z_s$, $s \in M'$. Then (54) has the unique continuous solution

$$z_s = h^s(t, y_s), \qquad s = 1,\dots, T, \tag{55}$$

$$h^s(t, y_s) \equiv t y_s, \qquad s \in M', \tag{56}$$

through $(1, y)$, where y is any arbitrary point in Z, and defined on a maximal t-interval $J(y)$.

Let $y = (y_1, \dots, y_s, \dots, y_T)$.

$$h(t, y) = (h^1(t, y_1), \dots, h^s(t, y_s), \dots, h^T(t, y_T)). \tag{57}$$

DEFINITION 2. A real-valued function F defined on an open subset Z^1 of R^T is called M-conditionally homogeneous of degree m or, for short $C_M H_m$, if

$$\begin{gathered} Z^1 \subset Z, \\ F(h(t, y)) = t^m F(y), \end{gathered} \tag{58}$$

for every $y \in Z^1$ and every $t \in J(y)$ for which $h(t, y) \in Z^1$.

The above definition collapses to the usual concept of homogeneity if M is empty.

We now present the generalization of Euler's Theorem. For this purpose, let $J^1(y)$, $y \in Z^1$, be the maximal interval containing 1 such that $t \in J^1(y)$ implies $h(t, y) \in Z^1$. Obviously, $J^1(y) \subset J(y)$.

THEOREM 2. *If $F: Z^1 \to R$ has a differential at $y \in Z^1$ and F is $C_M H_m$, then*

$$\sum_{s \in M} a^s(y_s) F_s(y) + \sum_{s \in M'} y_s F_s(y) = mF(y). \tag{59}$$

Proof. By hypothesis,

$$F(h(t, y)) = t^m F(y), \qquad t \in J^1(y). \tag{60}$$

There are two mutually exclusive cases to consider.

Case (i).

$$a_s(y_s) = 0, \qquad s = 1, \dots, T, \tag{61}$$

therefore

$$h(t, y) = y, \qquad t \in J(y). \tag{62}$$

Taking $t \in J(y) - \{1\}$, and using (61)–(62) in (60), we get

$$(t^m - 1) F(y) = 0,$$

implying either $m = 0$ or $F(y) = 0$ or both. This implies (59).

Case (ii).

$$\left.\frac{dz_s}{dt}\right|_{t=1} = a^s(y_s) \neq 0 \qquad \text{for some} \quad s \in \{1,\dots, T\}. \tag{63}$$

On the other hand,

$$t\,\frac{dz_s}{dt} = a^s(h^s(t, y_s)), \qquad t \in J(y). \tag{64}$$

Since a^s is continuous in z_s and $z_s \equiv h^s(t, y_s)$ is continuous in t, (63)–(64) imply that z_s is strictly monotone in t for t lying in some interval $J' \subset J^1(y)$ and $1 \in J'$. Hence

$$z_s \equiv h^s(t, y_s) \neq h^s(1, y_s) \equiv y_s, \qquad t \in J' - \{1\} \equiv J''$$

therefore

$$\|h(t, y) - y\| \neq 0 \qquad \text{for} \quad t \in J''.$$

Define, for $t \in J''$,

$$H(t, y) = \frac{F(h(t, y)) - F(y) - \nabla F(y) \cdot (h(t, y) - y)}{\|h(t, y) - y\|}. \tag{65}$$

Since F has a differential at y, it follows that

$$\lim_{t \to 1} H(t, y) = 0 \tag{66}$$

From (65), we have, for $t \in J''$,

$$\frac{F(h(t, y)) - F(y)}{t - 1} = H(t, y) \left\| \frac{h(t, y) - y}{t - 1} \right\| + \nabla F(y) \cdot \left(\frac{h(t, y) - y}{t - 1} \right). \tag{67}$$

By (54)–(56),

$$\lim_{t \to 1} \frac{h_r(t, y_r) - y_r}{t - 1} = a^r(y_r), \qquad r = 1,\dots, T. \tag{68}$$

Using (66) and (68) in (67), we have

$$\lim_{t \to 1} \frac{F(h(t, y)) - F(y)}{t - 1} = \sum_{r=1}^{T} a^r(y_r)\, F_r(y). \tag{69}$$

From (60),

$$\lim_{t \to 1} \frac{F(h(t, y)) - F(y)}{t - 1} = mF(y). \tag{70}$$

Equations (69)–(70) imply (59).

The converse of the generalized Euler's Theorem now follows.

THEOREM 3. *If $F: Z^1 \to R$ is a solution of the partial differential equation*

$$\sum_{s \in M} a^s(z_s) F_s(z) + \sum_{s \in M'} z_s F_s(z) = mF(z), \tag{71}$$

then $F(h(t, y)) = t^m F(y)$, for all $y \in Z^1$ and all $t \in J^1(y)$.

Proof. For $t \in J^1(y)$,

$$\begin{aligned} t \frac{\partial}{\partial t} F(h(t, y)) &= \sum_{s=1}^{T} a^s(h^s(t, y_s)) F_s(h(t, y)) \\ &= mF(h(t, y)) \qquad [\text{by } (71)] \\ &\Rightarrow F(h(t, y)) = t^m F(h(1, y)) = t^m F(y). \end{aligned}$$

THEOREM 4. *If $F: Z^1 \to R$ is $C_M H_m$, then F_s, $s \in M'$ is $C_M H_{m-1}$ on Z^1.*

Proof. Differentiate (58) withrespect to y_s, $s \in M'$, and use (56). Q.E.D.

Let us now illustrate the concept. Consider the function

$$F(z) = \left(\sum_{s \in M} f^s(z_s) \right) H^1(z') + H^2(z'),$$

where

(a) z' is the subvector of elements z_s, $s \in M'$,

(b) f^s is strictly monotone and continuously differentiable,

(c) H^1 and H^2 are homogeneous of degree 0 and m, respectively.

Take $a^s(z^s) = mf^s(z_s)/f^{s\prime}(z_s)$.

In this case, the solution path (55)–(56) takes the form

$$\begin{aligned} f_s(z_s) &= f_s(y_s)\, t^m, \qquad & s \in M, \\ z_s &= t y_s, \qquad & s \in M'. \end{aligned} \tag{72}$$

With z lying on the path (72), we have

$$F(z) = t^m \left[\sum_{s \in M} f^s(y_s) \right] H^1(ty') + H^2(ty'),$$
$$= t^m F(y), \qquad t \in J(y),$$

therefore F is $C_M H_m$.

Remark 5. Definition 1 given in the text can now be restated in the light of Theorem 3 with $M = \{s\}$, $m = 0$ and $a^s(z_s) = -\xi^s(z_s)$.

The direct utility function $u(x) = F(u^1(x^1),..., u^T(x^T))$, $x \in X$, is $C_s H_0$-in $z \equiv (u^1(x^1),..., u^T(x^T))$ if

$$F(h^s(t, u^s(x^s)), (tu^r(x^r))_{r \neq s}) = F(u^1(x^1),..., u^T(x^T))$$

for all $x \in X$ and all $t \in J(u^1(x^1),..., u^T(x^T))$.

This is synonymous with Definition 1 if F is of class C^1. Similarly, for the indirect utility function.

REFERENCES

1. C. Blackorby, D. Primont, and R. R. Russell, "Duality, Separability and Functional Structure: Theory and Economic Applications," North-Holland, Amsterdam, 1978.
2. W. M. Gorman, Separable utility and aggregation, *Econometrica* **27** (1959), 469–481.
3. W. M. Gorman, The structure of utility functions, *Rev. Econ. Studies* **35** (1968), 367–390.
4. W. Haque, "*S*-additive Preferences," (unpublished).
5. W. Haque, Nonsimultaneous additive separability, *Internat. Econ. Rev.* **15** (1974), 450–457.
6. L. J. Lau, Duality and the structure of utility functions, *J. Econ. Theory* **1** (1970), 374–396.
7. P. A. Samuelson, Constancy of the marginal utility of income, *in* "The Collected Scientific Papers of Paul A. Samuelson," Vol. I, (J. E. Stiglitz, Ed.), Chap. 5, M.I.T. Press, Cambridge, Mass., 1966.
8. P. A. Samuelson, Corrected formulation of direct and indirect additivity, *Econometrica* **37** (1969), 355–359.
9. M. Sono, The effect of price changes on the demand and supply of separable goods, *Internat. Econ. Rev.* **2** (1961), 239–271.
10. R. H. Strotz, The empirical implications of a utility tree, *Econometrica* **25** (1957), 269–280.

JOURNAL OF ECONOMIC THEORY 25, 255-268 (1981)

Majority Committees

PETER C. FISHBURN

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey 07974

Received January 20, 1980

Suppose an m-member committee is to be selected by a set of voters from a set X of $M > m$ nominees. If A is an m-member committee, and if as many voters prefer A to B as prefer B to A for every other committee B of m nominees, then A is a *majority committee of size m*. Although the existence of majority nominees ($m = 1$) has been extensively analyzed, little attention has been given to conditions that imply the existence of majority committees of size $m \geqslant 2$. Known restrictions on voters' preferences on X that guarntee the existence of a majority nominee could be applied directly to voters' preferences on m-member committees. However, this definitional exercise lacks intuitive appeal, and a different approach is taken in this paper. The paper presumes that profiles of voters' preferences on X are either dichotomous or single peaked. Both cases ensure the existence of a majority nominee. Independence–monotonicity assumptions are then used to connect voters' preferences on X to their preferences on committees of equal size. Although these assumptions guarantee the existence of majority committees when $m = 1$ and when $m = M - 1$, they do not generally do so when $1 < m < M - 1$. The latter observation motivates additional restrictions on profiles. In the dichotomous case, we consider profiles in which all voters have the same number k of nominees in their preferred subsets, and show that this restriction guarantees the existence of a majority committee of size m for $1 < m < M - 1$ only when $k = 1$ or $k = M - 1$. In the single-peaked case, we consider profiles in which all voters have the same most-preferred (peak) nominee, and prove that this guarantees the existence of a majority committee of size m for every m between 1 and $M - 1$.

1. INTRODUCTION

Let X be a finite set of $M \geqslant 3$ candidates or nominees, and let $\{1,..., n\}$ index a finite set of n individuals or voters, each of whom has a preference order on X. With respect to these preferences, X contains a *majority candidate* x if as many voters prefer x to y as prefer y to x for each $y \in X$. A restriction on individuals' preferences on X *guarantees* a majority candidate if, for every finite $n \geqslant 1$, X has a majority candidate whenever all individuals' preferences satisfy that restriction. Guaranteeing restrictions include single-peakedness, dichotomous preferences, value restriction, extremal restriction, and limited agreement [1, 4, 7, 10, 11]. Discussions of these restrictions have been oriented towards the election or social choice of one of the M candidates.

0022-0531/81/050255-14$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved

Since many social choice situations involve the selection of a committee of fixed size m from the M nominees, I shall consider here the existence of majority committees for each $m \in \{1, 2, \ldots, M-1\}$, where $m = 1$ is included for completeness. Let $X_m = \{A : A \subseteq X \text{ and } |A| = m\}$, the set of potential m-member committees. With respect to m and a set of n individuals, each of whom is presumed to have a preference order on X_m, $A \in X_m$ is a *majority committee of size m* (MC_m for short) if as many individuals prefer A to B as prefer B to A for every $B \in X_m$. Such an A is a *strong* MC_m if more individuals prefer A to B than prefer B to A for every $B \neq A$ in X_m. As with X, or X_1, a restriction or set of conditions on individuals' preferences on X_m *guarantees* a MC_m if, for every finite $n \geqslant 1$, there is a MC_m whenever all individuals' preferences on X_m satisfy the given restriction or set of conditions.

Interest in majority committees stems from Condorcet's principle [3], which, broadly interpreted, asserts that a majority alternative ought to be elected when one exists. Condorcet's principle, applied to a choice from X_m, seems reasonable in many committee-selection settings. However, some situations may suggest committee-selection criteria—such as notions of representativeness [2, 5, 6, 8, 9]—which could clash with that principle. Although I shall not discuss compatibility among different criteria, this matter clearly deserves further consideration.

Restrictions mentioned above for X could be applied to individuals' preferences on X_m to obtain conditions that guarantee a MC_m. However, this definitional exercise has questionable intuitive merit when $m \geqslant 2$, and therefore a different approach will be used. In particular, I shall examine cases in which individuals' preferences on X are either dichotomous or single peaked, and use independence–monotonicity assumptions to relate preferences on X_m for $m \geqslant 2$ to preferences on X. For example, if x, y and z are distinct nominees, and if an individual prefers x to y, then I shall assume that he prefers the two-member committee $\{x, z\}$ to the two-member committee $\{y, z\}$. Without such assumptions, which are clearly very strong, it would be futile to consider the existence of majority committees for $m \geqslant 2$ on the basis of preferences on the nominees in X.

The conclusions for guaranteed MC_m's that follow from the independence–monotonicity approach are mixed. On the positive side, we already know that dichotomous profiles and single-peaked profiles guarantee a MC_1. Subject to a simple independence–monotonicity condition, it will be shown that such profiles also guarantee a MC_{M-1}. However, when $1 < m < M-1$, it is usually possible to construct a profile of voters' preferences of the indicated type that has no MC_m. Because of this, we consider additional restrictions on dichotomous and single-peaked profiles, and use stronger independence–monotonicity conditions to see if these will guarantee MC_m's when $1 < m < M-1$. Under such restrictions and conditions we prove that

MC_m's always exist in the dichotomous case when all voters have exactly one preferred nominee, or all voters have exactly $M-1$ preferred nominees; and if all voters in a single-peaked profile have linear preferences and the same most-preferred nominee, then there is a MC_m.

The next section specifies our basic definitions and notation, introduces the first independence–monotonicity condition, and presents preliminary results that will guide our ensuing analysis. Sections 3 and 4 respectively deal further with dichotomous profiles and single-peaked profiles when $1 < m < M-1$. The paper concludes with a summary of the primary findings.

2. DEFINITIONS AND PRELIMINARY RESULTS

After introducing several basic notions that are needed throughout the paper, this section presents an independence–monotonicity condition that can be used to identify the members of a MC_m, if one exists. A simple example then shows that an MC_m might exist when there is no majority candidate. We then define the most general versions of dichotomous profiles and single-peaked profiles that we shall consider, and discuss known aspects of simple majorities for such profiles. The section concludes by observing that our basic independence condition guarantees a MC_{M-1} for all dichotomous and single-peaked profiles.

Three types of asymmetric order relations will be used. We shall say that a binary relation β on a set Y is (i) a *strict partial order* if it is asymmetric $[x\beta y \Rightarrow \text{not } (y\beta x)]$ and transitive $[x\beta y \;\&\; y\beta z \Rightarrow x\beta z]$, (ii) an *asymmetric weak order* if it is asymmetric and negatively transitive $[x\beta z \Rightarrow x\beta y \text{ or } y\beta z]$, and (iii) a *linear order* if it is a complete $[x \neq y \Rightarrow x\beta y \text{ or } y\beta x]$ asymmetric weak order.

A *profile* is a nonempty finite list $D = (\succ_1, \ldots, \succ_n)$ of strict partial orders on X. One can think of $\succ_i$ as the strict preference relation of individual i on the nominees in X. Some cases considered later assume that all $\succ_i$ are asymmetric weak orders or that all $\succ_i$ are linear orders. The strict, nonstrict and tied *simple majority relations* for a profile $D = (\succ_1, \ldots, \succ_n)$ are denoted respectively as $P(D)$, $R(D)$ and $I(D)$, where, for all $x, y \in X$,

$$\begin{aligned} xP(D)y \quad &\text{iff} \quad |\{i: x \succ_i y\}| > |\{i: y \succ_i x\}|, \\ xR(D)y \quad &\text{iff} \quad |\{i: x \succ_i y\}| \geqslant |\{i: y \succ_i x\}|, \\ xI(D)y \quad &\text{iff} \quad |\{i: x \succ_i y\}| = |\{i: y \succ_i x\}|. \end{aligned}$$

Thus, $P(D)$ and $I(D)$ are respectively the asymmetric and symmetric parts of $R(D) = P(D) \cup I(D)$.

Without its subscript i, $\succ$ denotes a generic preference order for an individual or voter, and $P_m(\succ)$ denotes this voter's asymmetric preference relation on X_m. We assume of course that $x \succ y$ iff $\{x\} P_1(\succ)\{y\}$. Although voters with the same $\succ$ on X might have different preference relations on X_m for $m \geqslant 2$, we shall assume that the following independence–monotonicity condition holds for all voters within all profiles.

CONDITION 1. *For all* $M \geqslant 3$, *all* $m \in \{1,..., M-1\}$, *and all* $A, B \in X_m$: *if* $|C| = m-1$, $A = C \cup \{x\}$ *and* $B = C \cup \{y\}$, *then* $AP_m(\succ)B$ *iff* $x \succ y$.

Thus, if two m-member committees are alike except that one contains x but not y, ad the other contains y but not x, then we assume that an individual prefers the first to the second if and only if he prefers x to y. Although Condition 1 is a simple (but not empirically vacuous) assumption, it can be very useful in attempting to locate a MC_m, if one exists.

LEMMA 1. *Suppose D is a profile, Condition* 1 *holds, and* $A \in X_m$. *Then A can be a* MC_m *only if there are no* $y \in X \setminus A$ *and* $x \in A$ *such that* $yP(D)x$, *and A can be a strong* MC_m *only if there are no* $y \in X \setminus A$ *and* $x \in A$ *for which* $yR(D)x$.

Proof. Under the lemma's hypotheses, suppose $x \in A$, $y \in X \setminus A$, and let $B = (A \cup \{y\}) \setminus \{x\}$. If $yP(D)x$, then more voters prefer y to x than x to y; hence, by Condition 1, more voters prefer B to A than A to B, and therefore A cannot be a MC_m. If $yI(D)x$, then Condition 1 implies that as many voters prefer B to A as A to B, so that A cannot be a strong MC_m. Q.E.D.

Suppose $M = 4$ and D is the three-voter profile $(xyzw, zxyw, yzxw)$, where $xyzw$ is a linear order with x most preferred and w least preferred. Since $P(D)$ is cyclic on $\{x, y, z\}$, there is no majority candidate, and Lemma 1 implies that there is no MC_2. If there is a MC_3, Lemma 1 says that it must be $\{x, y, z\}$, and it is easily seen that Condition 1 implies that this three-member committee is a strong MC_3. Henceforth, we shall consider only profiles that guarantee a majority candidate.

Dichotomous Profiles

DEFINITION 1. Let $\sim_i$ be the symmetric complement of $\succ_i$ on X, so that $x \sim_i y$ iff neither $x \succ_i y$ nor $y \succ_i x$. Profile $D = (\succ_1,..., \succ_n)$ is *dichotomous* if and only if, for each $i = 1,..., n$, there is a nonempty proper subset H_i of X such that, for all $x, x' \in H_i$ and all $y, y' \in X \setminus H_i$, $x \sim_i x'$, $y \sim_i y'$, and $x \succ_i y$.

Thus, D is dichotomous when every voter's $\succ_i$ is an asymmetric weak order that partitions X into two nonempty indifference classes, the more preferred of which is denoted H_i. For notational convenience, we sometimes express a dichotomous profile in terms of the H_i as

$$D = (H_1, H_2,..., H_n).$$

Given D in this form, it is easily seen (cf. [7]) that $P(D)$ on X is an asymmetric weak order with

$$xP(D)y \qquad \text{iff} \quad |\{i: x \in H_i\}| > |\{i: y \in H_i\}|.$$

Given Condition 1, Lemma 1 says that if D yields a MC_m, then it consists of m of the top nominees under $P(D)$, with the possibility of more than one MC_m if there are simple majority ties. Moreover, if D has a strong MC_m, say A, then $xP(D)y$ for all $x \in A$ and all $y \in X \backslash A$.

Single-Peaked Profiles

Our general definition of single-peaked profiles follows the characterization of Theorem 9.1 in [4]. The underlying linear order on X will be denoted $<_0$, with $x \leqslant_0 y$ iff $x <_0 y$ or $x = y$. Along with $\sim_i$ as the symmetric complement of $\succ_i$, we define $\succsim_i$ as the union of $\succ_i$ and $\sim_i$.

DEFINITION 2. Profile $D = (\succ_1, ..., \succ_n)$ is *single peaked* if and only if there is a linear order $<_0$ on X such that, for each $i = 1, ..., n$, $\succ_i$ on X is a strict partial order and there are unique $a_i, b_i \in X$ with $a_i \leqslant_0 b_i$ such that, for all $x, y, z \in X$,

(a) $y \succ_i x$ if either $x <_0 y \leqslant_0 a_i$ or $b_i \leqslant_0 y <_0 x$;

(b) $y \succsim_i x$ if $a_i \leqslant_0 y \leqslant_0 b_i$;

(c) $x \sim_i z$ if $x <_0 y <_0 z$, $x \sim_i y$ and $y \sim_i z$.

Part (a) of the definition says that voter i's preferences increase up to a_i and decrease after b_i under the order $<_0$. Part (b) requires indifference between each two nominees in the voter's "indifference plateau" from a_i to b_i, but it also allows something in this plateau to be indifferent to something outside the plateau interval. Part (c) requires the voter's indifference relation on X to be partly transitive, but it does not imply full transitivity of $\sim_i$, so that $\succ_i$ need not be an asymmetric weak order.

Part (1) of Theorem 9.2 in [4] says that, when D is single peaked, $P(D)$ on X is a strict partial order. Moreover, given (a_i, b_i) as in Definition 2 for $i = 1, ..., n$, Theorem 9.4 and its proof [4] assert that, if $c_1, c_2, ..., c_{2n}$ is a rearrangement of the sequence $a_1, ..., a_n, b_1, ..., b_n$ such that $c_1 \leqslant_0 c_2 \leqslant_0 \cdots \leqslant_0 c_{2n}$, then

$$yP(D)x \qquad \text{if} \quad x <_0 y \leqslant_0 c_n \quad \text{or} \quad c_{n+1} \leqslant_0 y <_0 x.$$

$$yR(D)x \qquad \text{if} \quad c_n \leqslant_0 y \leqslant_0 c_{n+1}.$$

Let $K = \{x: c_n \leqslant_0 x \leqslant_0 c_{n+1}\}$. Then K is the set of majority candidates, and $\{c_n\}$ is a strong MC_1 iff $c_n = c_{n+1}$. Under the hypotheses of Lemma 1, if A is a MC_m then $\{x \in A, x <_0 c_n\} \Rightarrow \{y: x \leqslant_0 y \leqslant_0 c_n\} \subseteq A$, and $\{x \in A, c_{n+1} <_0 x\} \Rightarrow \{y: c_{n+1} \leqslant_0 y \leqslant_0 x\} \subseteq A$. Moreover, if A is a strong MC_m, then the preceding $P(D)$ and $R(D)$ relationships along with Lemma 1 require $|K| \leqslant m$ with A a set of contiguous elements under $<_0$ that includes K.

Guarantees

Since $P(D)$ is transitive for all profiles D that are either dichotomous or single peaked, these restrictions on profiles guarantee a MC_1. Given Condition 1, they also guarantee a MC_{M-1}. This follows immediately from the ensuing theorem. For notational convenience we shall let $P^*(D)$ be the strict simple majority relation on X_{M-1}, so that when $D = (>_1, \ldots, >_n)$ and $A, B \in X_{M-1}$,

$$AP^*(D)\,B \qquad \text{iff} \quad |\{i: AP_{M-1}(>_i)\,B\}| > |\{i: BP_{M-1}(>_i)\,A\}|.$$

THEOREM 1. *Suppose D is a profile and Condition* 1 *holds. Suppose further that* $A, B \in X_{M-1}$ *with* $A = X \backslash \{x\}$ *and* $B = X \backslash \{y\}$. *Then* $AP^*(D)\,B$ *iff* $yP(D)\,x$. *Consequently, if* $P(D)$ *is a strict partial order on* X, *then* $P^*(D)$ *is a strict partial order on* X_{M-1}.

Proof. Under Condition 1 with $A = X \backslash \{x\}$ and $B = X \backslash \{y\}$, $AP_{M-1}(>_i)\,B$ iff $y >_i x$, and $BP_{M-1}(>_i)\,A$ iff $x >_i y$, so that $AP^*(D)\,B$ iff $yP(D)\,x$. The final assertion of the theorem then follows easily from this observation.

Q.E.D.

3. DICHOTOMOUS PROFILES AND $1 < m < M-1$

In place of Condition 1, our ensuing discussion of dichotomous profiles will use the stronger assumption that every "dichotomous voter" prefers $A \in X_m$ to $B \in X_m$ whenever A contains more of his preferred nominees than does B.

CONDITION 2. *For all* $M \geqslant 3$, *all* $m \in \{1, \ldots, M-1\}$, *and all* $A, B \in X_m$: *if* $>$ *on* X *is dichotomous with preferred subset* H (cf. Definition 1), *then* $AP_m(>)\,B$ *iff* $|A \cap H| > |B \cap H|$.

Since dichotomous profiles guarantee a MC_1 and a MC_{M-1} under Condition 1, or the stronger Condition 2, we shall consider only $1 < m < M-1$ in the rest of this section. Our first theorem for intermediate m is wholly negative in the guarantee sense.

THEOREM 2. *If* $1 < m < M-1$ *and Condition* 2 *holds, then there is a dichotomous profile that has no* MC_m.

Proof. Suppose $1 < m < M - 1$ and Condition 2 holds. Let $X = \{x_1, x_2, \ldots, x_M\}$ *and let* $D = (H_1, \ldots, H_7)$ with

$$H_i = \{x_1, \ldots, x_m\} \quad \text{for} \quad i = 1, 2, 3,$$

$$H_i = \{x_{m+1}\} \quad \text{for} \quad i = 4, 5,$$

$$H_i = \{x_{m+2}\} \quad \text{for} \quad i = 6, 7.$$

Lemma 1 and the discussion after Definition 1 imply that only $\{x_1, \ldots, x_m\}$ can be a MC_m. However, Condition 2 implies $\{x_3, x_4, \ldots, x_{m+1}, x_{m+2}\} P_m(>_i)\{x_1, \ldots, x_m\}$ for $i = 4, 5, 6$, and 7; hence there is no MC_m. Q.E.D.

Uniformly Dichotomous Profiles

In view of Theorem 2, we consider a severe restriction on dichotomous profiles that might yield guarantees for some case in which $1 < m < M - 1$. In particular, we shall examine situations in wich all H_i contain the same number of nominees.

DEFINITION 3. A dichotomous profile $D = (H_1, \ldots, H_n)$ is *k-uniform* if and only if $|H_i| = k$ for $i = 1, \ldots, n$.

The following theorem shows that uniformly dichotomous profiles guarantee a MC_m only in several special cases. An alternative form of the theorem says that, given $1 < m < M - 1$, $k \in \{1, 2, \ldots, M - 1\}$ and Condition 2, it is possible to construct a k-uniform dichotomous profile that has no MC_m iff $1 < k < M - 1$ and $M \geqslant 5$.

THEOREM 3. *Suppose* $1 < m < M - 1$, $k \in \{1, 2, \ldots, M - 1\}$, *and Condition 2 holds. Then k-uniform dichotomous profiles guarantee a* MC_m *if, and only if, either* $k = 1$, $k = M - 1$, *or* $(k, m, M) = (2, 2, 4)$.

Proof. We assume throughout that $1 < m < M - 1$, $k \in \{1, 2, \ldots, M - 1\}$, and that Condition 2 holds. The guarantee cases will be considered first. We then construct examples that have no MC_m when $1 < k < M - 1$ and $M \geqslant 5$.

Suppose first that dichotomous D is 1-uniform. Given $X = \{x_1, \ldots, x_M\}$, suppose for definiteness that $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_M$, where n_j is the number of voters whose preferred nominee is x_j. Let $A = \{x_1, \ldots, x_m\}$. Since the sum of no other m of the n_j can exceed $n_1 + \cdots + n_m$, it follows from Condition 2 that A is a MC_m. To elaborate, suppose $B \in X_m$ and $B \neq A$. All voters with preferred nominee in $A \cap B$ or in neither A nor B are indifferent between A and B. Those whose preferred nominee is in $A \backslash B$ have $A P_m(>_i) B$, and those whose preferred nominee is in $B \backslash A$ have $B P_m(>_i) A$. Since $|A| = |B|$, $|A \backslash B| = |B \backslash A|$; and since $n_j \geqslant n_k$ when $x_j \in A \backslash B$ and $x_k \in B \backslash A$, it follows that at least as many voters prefer A to B as prefer B to A.

Suppose next that D is $(M-1)$-uniform and with $X=\{x_1,\ldots,x_M\}$, that $n'_1 \leqslant n'_2 \leqslant \cdots \leqslant n'_M$, where n'_j is the number of voters who *omit* x_j from their set of $M-1$ preferred nominees. Again, with $A=\{x_1,\ldots,x_m\}$, it follows easily from Condition 2 that A is a MC_m.

To complete the guarantee cases, let $(k,m,M)=(2,2,4)$. Let $X=\{x_1,x_2,x_3,x_4\}$ and let n_{ij} for $1\leqslant i<j\leqslant 4$ be the number of voters in a 2-uniform dichotomous profile D whose preferred subset is $\{x_i,x_j\}$. Since $P(D)$ on X is an asymmetric weak order, assume for definiteness that $x_i R(D) x_j$ whenever $i<j$, so that

$$n_{12}+n_{13}+n_{14} \geqslant n_{12}+n_{23}+n_{24} \geqslant n_{13}+n_{23}+n_{34} \geqslant n_{14}+n_{24}+n_{34}.$$

The sum of the first two n_{ij} sums compared to the sum of the last two n_{ij} sums shows that $n_{12} \geqslant n_{34}$. Therefore, with $m=2$, $\{x_3,x_4\}$ cannot have a strict majority over $\{x_1,x_2\}$, since otherwise Condition 2 would require $n_{34}>n_{12}$. Suppose $\{x_1,x_3\}$ has a strict majority over $\{x_1,x_2\}$ which, by Condition 2, would require $n_{13}+n_{34}>n_{12}+n_{24}$, or $n_{13}+n_{23}+n_{34}>n_{12}+n_{23}+n_{24}$. But this contradicts $x_2 R(D) x_3$. Hence $\{x_1,x_2\}$ can't lose to $\{x_1,x_3\}$ by majority comparison. A similar result obtains for each of $\{x_1,x_4\}$, $\{x_2,x_3\}$ and $\{x_2,x_4\}$ in place of $\{x_1,x_3\}$, and therefore $\{x_1,x_2\}$ is a MC_2.

It remains to show that, whenever $M\geqslant 5$, $2\leqslant k\leqslant M-2$ and $2\leqslant m\leqslant M-2$, there is a k-uniform dichotomous profile that has no MC_m. We shall divide this into three main cases, according to whether $k=m$, $k>m$, or $k<m$. With $X=\{x_1,\ldots,x_M\}$, all cases will take

$$A=\{x_1,\ldots,x_m\},$$
$$B=\{x_1,\ldots,x_{m+2}\}.$$

Our profiles will be constructed so that, by Lemma 1, only A can be a MC_m, but in each case more voters will prefer B to A than prefer A to B under Condition 2, so that there will be no MC_m. Profiles will be listed by giving the number of voters whose preferred set (H_i) is shown to the right of the number, and we shall let n_j be the number of voters whose preferred sets contain x_j. For example, with $k=2$ and $m\geqslant 6$,

3: $\{x_1,x_2\}$,

2: $\{x_3,x_5\}$,

2: $\{x_4,x_6\}$,

is a 2-uniform profile in which three voters have $H_i=\{x_1,x_2\}$, two have $H_i=\{x_3,x_5\}$, and two have $H_i=\{x_4,x_6\}$. In this case, $(n_1,\ldots,n_6)=$

$(3, 3, 2, 2, 2, 2)$, and if $m = 2$—so that $A = \{x_1, x_2\}$ and $B = \{x_3, x_4\}$—then only A can be a MC_2, whereas B beats A by a 4-to-3 majority.

Case I. $2 \leqslant k = m \leqslant M - 2$. Suppose first that $m = M - 2$, so that $A = \{x_1, \dots, x_{M-2}\}$ and $B = \{x_3, \dots, x_M\}$, and consider the 11-voter $(M - 2)$-uniform profile

$$\begin{aligned}
&3: A,\\
&2: \{x_1\} \cup \{x_4, \dots, x_M\},\\
&2: \{x_2\} \cup \{x_4, \dots, x_M\},\\
&2: \{x_1\} \cup \{x_3, \dots, x_{M-2}\} \cup \{x_{M-1}\},\\
&2: \{x_2\} \cup \{x_3, \dots, x_{M-2}\} \cup \{x_M\}.
\end{aligned}$$

Since $M \geqslant 5$, $\{x_3, \dots, x_{M-2}\}$ is not empty, and the four voters in the second and third lines prefer B to A since only one of the B nominees (x_3) is missing from their H_i's, whereas two of the A nominees (x_2 and x_3, or x_1 and x_3) are not in their H_i's. By Condition 2, the final four voters are indifferent between A and B. Therefore B has a 4-to-3 majority over A. But since $n_1 = n_2 = n_3 = 7$, $n_4 = \cdots = n_{M-2} = 11$, and $n_{M-1} = n_M = 6$, only A could be a MC_{M-2}, and hence there is no MC_{M-2}.

To complete Case I, suppose in this paragraph that $k = m < M - 2$, and consider the profile

$$\begin{aligned}
&3: A,\\
&2: [B \backslash \{x_{m+1}\}] \cup \{x_M\},\\
&2: [B \backslash \{x_{m+2}\}] \cup \{x_M\},\\
&2: \{x_1\} \cup [B \backslash \{x_{m+1}\}],\\
&2: \{x_2\} \cup [B \backslash \{x_{m+2}\}].
\end{aligned}$$

Again, B has a 4-to-3 majority over A (the last four voters are indifferent between the two). Moreover, $n_1 = n_2 = 5$, $n_3 = \cdots n_m = 11$ (unless $m = 2$), $n_{m+1} = n_{m+2} = n_M = 4$, and all other $n_j = 0$, so that only A could be a MC_m.

Case II. $2 \leqslant m < k \leqslant M - 2$. When $k = m + 1$, we use the profile

$$\begin{aligned}
&3: A \cup \{x_{m+3}\},\\
&2: \{x_1\} \cup B,\\
&2: \{x_2\} \cup B.
\end{aligned}$$

Since $m + 1 \leqslant M - 2$, $m + 3 \leqslant M$, so that $x_{m+3} \in X$. Here B has a 4-to-3 majority over A; $n_1 = n_2 = 5$, $n_3 = \cdots = n_m = 7$, $n_{m+1} = n_{m+2} = 4$, $n_{m+3} = 3$, and $n_j = 0$ for $j > m+3$. Assume henceforth for Case II that $k \geqslant m+2$. Consider a profile that begins with

$$3\colon A \cup \{x_{m+3},\dots, x_{k+2}\},$$

$$2\colon \{x_1\} \cup B \cup \{x_{m+3},\dots, x_{k+1}\},$$

$$2\colon \{x_2\} \cup B \cup \{x_{m+3},\dots, x_{k+1}\},$$

where B has a 4-to-3 majority over A. Now if $k = m+2$, add voters with $H_i = \{x_1,\dots, x_{m+2}\} = A \cup B$, and if $k > m+2$, add an equal number of voters with $H_i = A \cup B \cup C$ for each way that C with $|C| = k-(m+2)$ can be chosen from $\{x_{m+3},\dots, x_M\}$. All added voters, say N in number, are indifferent between A and B under Condition 2, and we have $n_1 + n_2 = N+5$, $n_3 = \cdots = n_m = N+7$, $n_{m+1} = n_{m+2} = N+4$, and $n_j < N+5$ for each $j \geqslant m+3$ when N is sufficiently large. Hence only A could be a MC_m when N is large, but since B has a strict majority over A, there is no MC_m.

Case III. $2 \leqslant k < m \leqslant M-2$. We begin our k-uniform profile with

$$3\colon \{x_1,\dots, x_k\},$$

$$2\colon \{x_3,\dots, x_{k+1}, x_{m+1}\},$$

$$2\colon \{x_3,\dots, x_{k+1}, x_{m+2}\},$$

for which B has a 4-to-3 majority over A. If $m = k+1$, we are finished, since then $n_1 = n_2 = 3$, $n_3 = \cdots = n_{m-1} = 7$ (if $k \geqslant 3$), $n_m = 4$, $n_{m+1} = n_{m+2} = 2$, and $n_j = 0$ for $j > m+2$. Assume henceforth that $m \geqslant k+2$, or $m-2 \geqslant k$, in which case $|A \cap B| = m-2 \geqslant k$. We then add voters with k-element preferred sets from $A \cap B = \{x_3,\dots, x_m\}$ in sufficient numbers so that all n_j for $j \leqslant m$ exceed 2. Hence only A can be a MC_m, and since all voters with $H_i \subseteq A \cap B$ are indifferent between A and B, B maintains its 4-to-3 majority over A. Q.E.D.

4. SINGLE-PEAKED PROFILES AND $1 < m < M-1$

Given two m-member committees A and B, and a preference order $\succ$ on X, we shall say that *A dominates B under $\succ$* if there is a mapping σ from A onto B such that either $x = \sigma(x)$ or $x \succ \sigma(x)$ for every $x \in A$, and $x \succ \sigma(x)$ for some $x \in A$. When $\succ$ is a strict partial order, it is easily proved that the dominance relation on X_m is also a strict partial order. In place of Condition 1, our ensuing discussion will use dominance according to the following condition.

CONDITION 3. *For all* $M \geqslant 3$, *all* $m \in \{1, \ldots, M-1\}$, *and all* $A, B \in X_m$: *if* A *dominates* B *under* $\succ$, *then* $AP_m(\succ)B$. *Moreover, if* A *does not dominate* B *under* $\succ$ *but* $x \succ y$ *for some* $x \in A$ *and all* $y \in B$, *then it may be true that* $AP_m(\succ)B$.

In this section, I shall use a restricted version of single-peakedness, partly to highlight the negative nature of the next theorem in so far as guarantees are concerned.

DEFINITION 4. Profile $D = (\succ_1, \ldots, \succ_n)$ is *strongly single peaked* if and only if it is single peaked and $\succ_i$ on X is a linear order for $i = 1, \ldots, n$.

In comparison with Definition 2, strong single-peakedness requires $a_i = b_i$ for all i, and no individual is indifferent between distinct nominees. The following theorem is a single-peaked companion to Theorem 2 for dichotomous profiles.

THEOREM 4. *If* $1 < m < M - 1$ *and Condition* 3 *holds, then there is a strongly single-peaked profile that has no* MC$_m$.

Proof. Let $X = \{x_1, \ldots, x_M\}$, and assume for definiteness that the underlying linear order $<_0$ on X has $x_1 <_0 x_2 <_0 \cdots <_0 x_M$. Given $1 < m < M - 1$, let D be a strongly single-peaked pofile with three voters who have the following linear preference orders on X:

voter 1: $x_1 x_2 \cdots x_M$ (x_1 most preferred),
voter 2: $x_M x_{M-1} \cdots x_1$ (x_M most preferred),
voter 3: $x_2 \cdots x_{m+1} x_1 x_{m+2} \cdots x_M$.

Since the first two voters' orders are self-dual, $P(D)$ is simply the linear order of voter 3. Since the first part of Condition 3 implies Condition 1, Lemma 1 allows us to conclude that only $A = \{x_2, \ldots, x_{m+1}\}$ can be a MC$_m$. Let $B = \{x_1, x_M\}$ if $m = 2$, and $B = \{x_1, x_2, \ldots, x_{m-1}, x_M\}$ if $m \geqslant 3$. Since voter 1 likes x_1 best, and voter 2 likes x_M best, the latter part of Condition 3 permits us to presume that both voters prefer B to A. Assuming this to be the case, B has a 2-to-1 majority over A and there is no MC$_m$. Q.E.D.

A Unanimous Peak

Theorem 3 showed that the imposition of uniformity on dichotomous profiles had only a minor effect on the results of Theorem 2. In contrast to this, we now consider a restriction on strongly single-peaked profiles that always guarantees a MC$_m$ under Condition 3. In particular, we shall suppose that all voters have the same most-preferred nominee. If $x_1 <_0 \cdots <_0 x_M$, and if x_k is the unanimously most-preferred nominee, then all voters have the same linear preference order $x_{k-1} \cdots x_1$ on the $k-1$ nominees to the left of x_k under $<_0$, and all voters have the same linear preference order $x_{k+1} \cdots x_M$ on the $M-k$ nominees to the right of x_k under $<_0$. Since these two orders can be merged in $\binom{M-1}{k-1}$ ways to give a linear order on $X \backslash \{x_k\}$, a unanimously-peaked strongly single-peaked profile can contain as many as $\binom{M-1}{k-1}$ different preference orders. For example, if $k=1$, then all voters have the same order $x_1 x_2 \cdots x_M$, but if $k=3$ with $M=5$, then each of $x_3x_2x_1x_4x_5$, $x_3x_2x_4x_1x_5$, $x_3x_2x_4x_5x_1$, $x_3x_4x_2x_1x_5$, $x_3x_4x_2x_5x_1$, and $x_3x_4x_5x_2x_1$ could appear in the profile.

THEOREM 5. *If* $1 < m < M-1$ *and Condition* 3 *holds, then every strongly single-peaked profile in which all voters have the same most-preferred nominee has a* MC_m.

Proof. Given $1 < m < M-1$ and Condition 3, assume that $D = (\succ_1, \ldots, \succ_n)$ is strongly single-peaked with $x_1 <_0 \cdots <_0 x_M$ and that x_k is every voter's most-preferred nominee. Since $P(D)$ on X is a strict partial order that includes the subset linear orders $x_k x_{k-1} \cdots x_1$ and $x_k x_{k+1} \cdots x_M$, it follows without difficulty that there is an m-element subset of successively adjacent nominees in X, say

$$A = \{x_a, x_{a+1}, \ldots, x_{a+m-1}\},$$

with $a \leqslant k \leqslant a+m-1$ such that:

- (*) if $a+m-1 < M$, then $x_a R(D) x_{a+m}$, so that at least half the voters prefer x_a to x_{a+m};
- (**) if $a > 1$, then $x_{a+m-1} R(D) x_{a-1}$, so that at least half the voters prefer x_{a+m-1} to x_{a-1}.

Such an A can be constructed by starting at $x_k \in A$ and moving outwards in both directions from x_k, using $P(D)$ to compare nominees on different sides of the peak.

Given A, assume for definiteness that $a < k$ (the proof for $a+m-1 > k$ is similar), and let $B \neq A$ be any other m-member committee in X_m. We shall prove that A dominates B under $>_i$ for at least half the $i \in \{1, \ldots, n\}$, so that A is a MC_m according to the first part of Condition 3. In applying dominance, we use $x = \sigma(x)$ for all $x \in A \cap B$, and henceforth will not bother with the elements in $A \cap B$. Let

t_1 be the number of nominees in $\{x_a, \ldots, x_{k-1}\} \backslash B$;
t_2 be the number of nominees in $\{x_k, \ldots, x_{a+m-1}\} \backslash B$;
s_1 be the number of nominees in $\{x_1, \ldots, x_{a-1}\} \cap B$;
s_2 be the number of nominees in $\{x_{a+m}, \ldots, x_M\} \cap B$.

Since $A, B \in X_m$ and $A \neq B$, $t_1 + t_2 = s_1 + s_2 \geqslant 1$. If $t_1 = s_1$, then everyone prefers A to B under Condition 3 by an obvious choice of σ. Suppose next that $t_1 > s_1$. If $s_1 > 0$, specify s_1 of the t_1 nominees, all of which are preferred by everyone to the s_1 nominees because of single-peakedness; moreover, at least half of the voters prefer the rest of the t_1 nominees to all those that make up s_2 according to (*), and these same voters prefer everything in t_2 to everything in s_2 according to single-peakedness. It then follows from Condition 3 that A dominates B for at least half the voters. Finally, if $t_1 < s_1$, then $t_2 > s_2$, and an argument like that just given, which uses (**) instead of (*), shows that A dominates B for at least half the voters. Q.E.D.

5. SUMMARY

When $1 \leqslant m \leqslant M - 1$, there is a majority committee of size m (MC_m), say A, within the set of M nominees, if as many voters prefer A to B as prefer B to A for every other committee B of m nominees. This paper examined the existence of MC_m's when voters' preferences on the M nominees form dichotomous profiles or single-peaked profiles, and when specified independence-monotonicity conditions hold between voters' preferences on the nominees and their preferences on equal-sized committees of nominees. Under the first of these conditions, Condition 1, it was noted that there is always a MC_1 and a MC_{M-1} for the dichotomous and single-peaked cases.

Using a second and stronger independence condition for dichotomous profiles, we proved that if $1 < m < M - 1$, then there is a dichotomous profile that has no MC_m (Theorem 2). However, if all voters have the same number k of nominees in their preferred subsets, then, when $1 < m < M - 1$, there is always a MC_m in the dichotomous case if $k = 1$ or $k = M - 1$ or $(k, m, M) = (2, 2, 4)$; but if $1 < k < M - 1$ and $M \geqslant 5$, then there is a k-uniform dichotomous profile that has no MC_m (Theorem 3).

Single-peaked profiles for $1 < m < M - 1$ were then examined further under a dominance condition for voters' preferences on m-member committees. Under a simplified version of single-peakedness in which all voters have linear preference orders on the nominees, we proved first that, when $1 < m < M - 1$, there is a single-peaked profile that has no MC_m (Theorem 4); however, if all voters have the same most-preferred nominee, then every single-peaked profile has a MC_m (Theorem 5).

REFERENCES

1. D. BLACK, "The Theory of Committees and Elections," Cambridge Univ. Press, Cambridge, 1958.
2. J. R. CHAMBERLAIN AND P. N. COURANT, "On the Selection of Representative Committees," Institute of Public Policy Studies Discussion Paper No. 108, The University of Michigan, 1977.
3. MARQUIS DE CONDORCET, "Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité de voix," Paris, 1785; reprinted, Chelsea, New York, 1973.
4. P. C. FISHBURN, "The Theory of Social Choice," Princeton Univ. Press, Princeton, N. J., 1973.
5. I. J. GOOD AND T. N. TIDEMAN, "A Principle for Selecting a Representative Committee," Virginia Polytechnic Institute and State University, 1976.
6. C. G. HOAG AND G. H. HALLETT, "Proportional Representation," MacMillan Co., New York, 1926.
7. K. INADA, A note on the simple majority decision rule, *Econometrica* **32** (1964), 525–531.
8. J. S. MILL, "Considerations on Representative Government," Harper and Brothers, New York, 1862.
9. H. F. PITKIN, Ed., "Representation," Atherton, New York, 1969.
10. A. K. SEN, "Collective Choice and Social Welfare," Holden-Day, San Francisco, 1970.
11. A. K. SEN AND P. K. PATTANAIK, Necessary and sufficient conditions for rational choice under majority decision, *J. Econ. Theory* **1** (1969), 178–202.

JOURNAL OF ECONOMIC THEORY 25, 269–282 (1981)

Fair Allocations in Large Economies*

PAUL CHAMPSAUR

Unité de Recherche, INSEE,
18 Bd Adolphe Pinard, 75675 Paris cedex 14, France

AND

GUY LAROQUE

ENSAE, 3, Avenue Pierre-Larousse, 92240 Malakoff, France

Received June 5, 1980; revised September 23, 1980

The purpose of this paper is the study of problems of equity in an economy with many agents. Conditions are given under which every equitable and Pareto optimal allocation is a competitive equilibrium with equal incomes for all agents. The results are also of potential relevance for the literature on incentive compatibility since the allocations that are obained through incentive mechanisms in large economies turn out to be equitable.

The purpose of this paper is the study of problems of equity in an economy with many agents. Our results are also of potential relevance for the literature on incentive compatibility since the allocations that are obtained through incentive mechanisms in large economies turn out to be equitable in our sense.

The intuition behind our results can be found in the works of Varian [1976] and Hammond [1979] which bear respectively on fairness and incentive compatibility. Our work aims at making their arguments precise, while establishing the relationship between the two approaches.

The plan of our paper is as follows. The first section is heuristic and tries to give the main ideas that stimulated our interest in the field. The second part describes precisely the model and the results. The proofs are gathered in the third and final section.

---

* We are indebted to B. Cornet and A. Monfort for helpful discussions. We are also grateful to an anonymous referee of this Journal for his remarks.

After the completion of this article, we were informed of the existence of recent works on the same subject by N. L. Kleinberg (Fair allocations and equal income, *J. Econ. Theory* **23** (1980), 189–200) and by A. Mac Lennan in his Ph.D. Thesis.

0022-0531/81/050269-14$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

1. GENERAL PRESENTATION

A. *Description of the Results*

An allocation is *equitable* if every agent in the economy prefers his consumption bundle to the consumption of other agents. More precisely, let A be the set of agents characteristics, $\mathbf{x}(a)$ be the consumption bundle of agent a, $\mathbf{x}(A)$ be the subset of R^l containing all the vectors $\mathbf{x}(a)$ for a in A. A necessary and sufficient condition for $\mathbf{x}$ to be equitable is that for all a, $\mathbf{x}(a)$ maximizes the utility of agent a on $\mathbf{x}(A)$.

We restrict ourselves to the study of *Pareto-optimal* allocations, that is allocations such that all the normals to the indifference hypersurfaces of the agents at $\mathbf{x}(a)$ are parallel to a vector p, which we call the price system.

An allocation is *fair* if it is both equitable and Pareto-optimal.

It is straightforward to check that a Pareto-optimal allocation which gives the same income $r = p \cdot \mathbf{x}(a)$ to all the agents, is equitable ($\mathbf{x}(A)$ is contained in the fixed budget hyperplane). Our results is that, if A the set of agents characteristics has nice topological properties (in particular connectedness) and if the references are smooth, then *every* fair allocation gives the same income to all the agents. The crucial assumption is here the connectedness of the set A: in this situation, it means that there is no ground from the point of view of equity for income transfers based on differences in the preferences of the agents, at least if one wants to preserve efficiency. Conversely, if there are disconnected classes of agents, there may exist fair allocations with different incomes (i.e., one can give more income to a paralytic for him to buy a wheelchair).

Remark 1. The notion of fairness that we use should be contrasted with that of coalitional fairness introduced by Schmeidler and Vind [1972] where the agents compared their own consumption with the one of the various coalitions. A related point is that since they worked by replicating finite economies, the set A is never connected in their paper, and thus the situation that they consider does not satisfy our assumptions.

Remark 2. Our result apply with minor changes to the study of fair net trades (where, given his initial endowment, every agent prefers his net trade to the net trade of any other agent). Under similar assumptions, one gets the conclusion that every fair net trade allocation is a competitive equilibrium.

Remark 3. The crucial assumption concerns the topological structure of the space of the *characteristics* of the agents belonging to the economy. For this set to be connected, one needs in general an economy with many agents. However, these agents may have very different sizes: the economy may have atoms.

B. *Intuition behind the Result*

A fair allocation is completely characterized by the following property: let $R(a, \mathbf{x}(a))$ be the set of vectors x strictly preferred by a to $\mathbf{x}(a)$. Then the intersection of $\mathbf{x}(A)$ with $R(a, \mathbf{x}(a))$ is empty and $p \cdot x > p \cdot \mathbf{x}(a)$ for all x in $R(a, \mathbf{x}(a))$. From this property, we want to show that, if A is connected, the set $\mathbf{x}(A)$ is contained in a fixed hyperplane $p \cdot x = \text{constant}$. The difficulty in the proof lies in the fact that we have to deduce from the topological properties of A a property of the set $\mathbf{x}(A)$ which is defined in a very round about way through the agents' preferences.

To give some intuition concerning the role of the various assumptions, one can try to decompose the proof in successive steps:

(1) $\mathbf{x}(a)$ is essentially bounded. To derive this property, we have to assume that the preferences of two close agents are not too different at infinity. This has lead us to introduce a finer topology on preference relations than the usual topology of closed convergence which does not control the behaviour of the preferences at infinity.

(2) $\mathbf{x}(a)$ is a continuous function of a. Knowing that $\mathbf{x}(a)$ is bounded, this is a simple consequence of the continuity and of the strict quasi-concavity of preferences. The graphical proof (Fig. 1) can be made precise along the following lines. Let (a^k) be a sequence of agents converging to a, and $\bar{x}$ an accumulation point of the sequence $\mathbf{x}(a^k)$. First remark that $\bar{x}$ is at least as good as $\mathbf{x}(a)$ for agent a: in fact if $\mathbf{x}(a)$ was strictly preferred by a to $\bar{x}$, then $\mathbf{x}(a)$ would be strictly preferred by a^k to $\mathbf{x}(a^k)$ for large enough k, and a^k would envy a, a contradiction. For similar reasons, $\bar{x}$ cannot be strictly preferred to $\mathbf{x}(a)$. Thus $\bar{x}$ is equivalent to $\mathbf{x}(a)$ for a. Then remark that $p \cdot R(a^k, \mathbf{x}(a^k)) > p \cdot \mathbf{x}(a^k)$ for all k and thus $p \cdot R(a, \bar{x}) \geqslant p \cdot \bar{x}$. Since

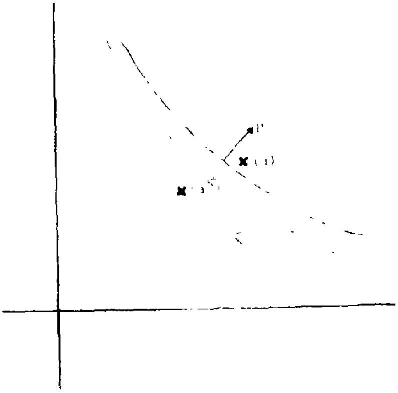


FIGURE 1

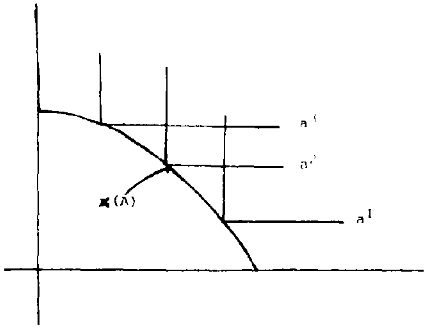


FIGURE 2

$R(a, \bar{x})$ is equal to $R(a, \mathbf{x}(a))$ and the preferences are strictly convex, it follows that $\bar{x} = \mathbf{x}(a)$. (This is essentially the argument given by Varian [1976]).

(3) Now, through the continuity of $\mathbf{x}$, the connectedness of A implies the connectedness of $\mathbf{x}(A)$. This is clearly not enough to get the desired results, as shown by Fig. 2, if preferences are not differentiable. On the other hand, if the utility functions are continuously differentiable, it is not possible to construct an $\mathbf{x}(A)$ as in Fig. 2: there is not much room left for $\mathbf{x}(A)$ and it looks as if $\mathbf{x}(A)$ has to be contained in a hyperplane $p \cdot x = \text{constant}$. However, the properties that we have obtained for $\mathbf{x}(A)$ do not seem to be sufficient to get the result we are looking for.[1] We have thus been lead to introduce smoothness assumptions on the utility functions to complete our proof.

The difficulty is linked with the differentiability of $\mathbf{x}$. In fact, if one assumes that $\mathbf{x}(a)$ is differentiable, the rest of the proof is straightforward. Let $u(a, x)$ be the utility function of agent a when he consumes x. Now, $u(a, x(a))$ is the maximum of $u(a, x(b)$ when b varies in A. Thus $D_x u \cdot dx/da = 0$, which implies $p \cdot dx/da = 0$. Then a strong property of connectedness of A ensures that $p \cdot \mathbf{x}(a)$ is constant. (This argument also was given by Varian [1976].)

The proof that we give in the third section below does not follow the lines suggested above because we have to tackle in full generality the fact that, in continuum economies, properties hold only almost everywhere: if one excludes a null set the topological properties of A are not preserved.

---

[1] It seems possible to build examples of economies with continuously differentiable preferences for which there exist continuous fair allocations which do not satisfy the equal income property. We are grateful to Professor R. Aumann for helpful discussions on this problem. Professor N. L. Kleinberg told us that such an example can be found in his Ph.D. Thesis.

C. *Relationship with the Incentive Compatible Literature*

Let us consider a *non atomic* economy and an *incentive procedure* which is defined as follows.

There is a set of messages M, the same for all agents. Every agent a send a message $\mathbf{m}(a)$ to the central planner and the planner chooses an allocation $\mathbf{x}$ as a function of $\mathbf{m}$. We shall say that the allocation mechanism is *symmetric* if it does not depend on the agents' name, i.e., given the distribution ν of the messages on the space M, the consumption bundle $\mathbf{x}(a)$ that agent a receives is a function $f(\mathbf{m}(a), \nu)$ which does not depend directly on a. Finally, we assume that each agent plays Nash.

Consider now an allocation obtained through an incentive procedure. Since a is small and cannot change ν by its own action, $\mathbf{x}(a)$ maximizes the utility of agent a on the set $f(M, \nu)$, which is independent of a. Since $\mathbf{x}(A)$ is contained in $f(M, \nu)$, $\mathbf{x}(a)$ maximizes the utility of agent a on $\mathbf{x}(A)$. Thus any allocation obtained through a *symmetric incentive procedure* is equitable.

Consequently, if an allocation obtained through a symmetric incentive procedure is Pareto-optimal, it is fair. Our result then says that, if the set of agents' characteristics is connected, the allocation gives equal incomes to everybody. To put it differently, there is no way to undertake lump sum transfers in such an economy, without loosing either Pareto-optimality or the possibility of getting information through incentives (or decentralizing the allocation mechanism).

2. RESULTS

We consider an economy with l commodities and a continuum of consumers (or traders).

All consumers have the same consumption set P, the interior of the positive orthant of $\mathscr{R}^l$: $P = \{x \text{ in } \mathscr{R}^l,\ x \gg 0\}$. The tastes of a consumer are characterized by a preference relation defined on P, which is assumed to be a complete, continuous and monotonic preordering. Under these assumptions, the preference relation can be represented by a unique utility function defined on P, with value in $\mathscr{R}_{++} = \{x \in \mathscr{R},\ x \gg 0\}$ such that

$$\text{for all } t \text{ in } \mathscr{R}_{++}, \qquad u(te) = t,$$

where e is the vector of P with all coordinates equal to 1 (see Kannaï [1970]). From now on, we shall systematically use the above "Kannaï representation" of the preference relations.[2]

---

[2] In fact this representation seems to date back to H. Wold (A synthesis of pure demand analysis, *Scand. Aktuar. T.* 1943 (26)).

We first assume that the utility function u represents a smooth preference relation in the sense of Debreu [1972]:

(A-1) *u is a twice continuously differentiable function from P into $\mathbb{R}_{++}$. It is strictly monotonic (the vector $Du(x)$ of first derivatives of u has everywhere all its components strictly positive), and strictly quasi concave. Moreover the indifference hypersurfaces $\{x \text{ in } P,\ u(x) = \text{constant}\}$ have everywhere a nonzero gaussian curvature, and their closure in $\mathbb{R}^l$ is contained in P.*

We shall consider a parametrized family of consumers' characteristics. The space E of parameters is an open subset of $\mathbb{R}^m$. The parametrization is described through a function $\mathbf{u}$ defined on $E \times P$, with values in $\mathbb{R}_+$. For every element a in E, the utility function $\mathbf{u}(a, \cdot)$ is the Kannaï representation of the preference relation corresponding to a, and satisfies assumption (A.1).

(A-2) *The function $\mathbf{u}(a, x)$ from $E \times P$ into $\mathbb{R}_+$ is continuous. The vector of its first derivatives with respect to x, $D\mathbf{u}(a, x)$, from $E \times P$ into $\mathbb{R}^l$, is locally lipschitzian.*

An economy is defined by a (Borel) probability measure μ on $\mathbb{R}^m$, with a support A contained in E, and a vector ω in P of aggregate resources.

An allocation is a measurable function $\mathbf{x}$ defined on A, with values in P. It is feasible if

$$\int_A \mathbf{x}\, d\mu = \omega.$$

An allocation $\mathbf{x}$ is *Pareto optimal* if it is feasible and if there exists no feasible allocation $\mathbf{y}$ such that

$$\mathbf{u}(a, \mathbf{y}(a)) > u(a, \mathbf{x}(a)), \qquad \mu \text{ almost everywhere in } A.$$

If $u(a, \mathbf{x}(b)) > u(a, \mathbf{x}(a))$, we say that agent a is envying agent b. Agent a is *envious* at the allocation $\mathbf{x}$ if the set of agents which he is envying has a strictly positive measure. This leads us to the following definition:

DEFINITION 1. Let $\mathbf{x}$ be an allocation, and

$$B(\mathbf{x}) = \{(a, b) \text{ in } A \times A,\ u(a, \mathbf{x}(b)) > \mathbf{u}(a, \mathbf{x}(a))\}.$$

The allocation $\mathbf{x}$ is equitable if

$$\mu \otimes \mu(B(\mathbf{x})) = 0,$$

where $\mu \otimes \mu$ is the product measure on $A \times A$.

By Fibini's theorem, Definition 1 amounts to say that an allocation is equitable when the set of envious agents is negligible.

Definition 2. An allocation is fair if it is both equitable and Pareto-optimal.

A price system p is an element of the unit simplex $S = \{p \text{ in } P, \|p\| = 1\}$.

Definition 3. An allocation $\mathbf{x}$ is an egalitarian competitive equilibrium if it is feasible and if there exists a price system p in S such that, μ almost everywhere in A:

$$\mathbf{x}(a) \text{ maximizes } \mathbf{u}(a, x) \text{ on } \{x \text{ in } P,\ p \cdot x = p \cdot \omega\}.$$

Under assumption A-1, it is well known that, if $\mathbf{x}$ is a Pareto optimal allocation, there exists a unique price system p sustaining this allocation: $\mathbf{x}(a)$ maximizes $\mathbf{u}(a, x)$ on $\{x \text{ in } P,\ p \cdot x = p \cdot \mathbf{x}(a)\}$, almost everywhere in A. Therefore, with every Pareto optimal allocation $\mathbf{x}$, we can associate an income distribution $\mathbf{r}$, a function defined on A with values in $\mathscr{R}_+$, by the relationship $\mathbf{r}$, a function defined on A with values in $\mathscr{R}_+$, by the relationship $\mathbf{r}(a) = p \cdot \mathbf{x}(a)$. Since, by the feasibility condition $\int_A \mathbf{r}\, d\mu = p \cdot \omega$, an egalitarian competitive equilibrium is simply a Pareto optimal allocation such that the corresponding income distribution is egalitarian, i.e., is equal to a constant function, except on a negligible subset of A.

It is easy to see that every egalitarian competitive equilibrium is fair. We are interested in conditions under which the converse is true. For this, we need additional topological properties on A.

(A-3) *A is the closure in E of an open and connected subset of E*(3).

Assumptions (A-1), (A-2) and (A-3) do not seem to imply the (essential) boundedness of a fair allocation even if, in addition, A is compact. More precisely, we say that the income distribution $\mathbf{r}$ is bounded if there exists $\bar{r}$ such that $\mathbf{r}(a) \leqslant \bar{r}$, μ almost everywhere in A.

Proposition 1. *Under Assumptions* (A-1), (A-2) *and* (A-3), *if* $\mathbf{x}$ *is a fair allocation such that the corresponding income distribution is bounded, then* $\mathbf{x}$ *is an egalitarian competitive equilibrium.*

In order to bound the fair allocations, we had to introduce another assumption which aims at controlling the behaviour of neighbour agents at infinity.

---

[3] All the results are still valid if we replace (A-3) by the following assumption (A-3′) A is compact and connected by locally lipschitzian arcs.

(A-4) *For all agents* a *in* E, *for all strictly positive* ε, *there exists* η *such that*

$$\|b-a\| < \eta \qquad \textit{implies} \qquad |\mathbf{u}(a,x) - \mathbf{u}(b,x)| < \varepsilon(1 + \|x\|).$$

Behind this assumption, there is the need for the parametrization of the preference relations to be continuous when the set of preferences is endowed with a topology[4] which is finer than the standard topology of closed convergence introduced in the works of Kannaï [1970] and Hildenbrand [1970] (while coarser than the topology of the Hausdorff distance of Debreu [1969]).

THEOREM. *Under assumptions* (A-1) *to* (A-4), *if* $\mathbf{x}$ *is a fair allocation, then* $\mathbf{x}$ *is an egalitarian competitive equilibrium.*

3. PROOF OF THE RESULTS

We begin by proving two lemmas which are derived from our parametrization of the preference relations.

For a given price system p in S and strictly positive income r, the demand function of an agent a in E, $\xi(a,p,r)$, is the vector of P which maximizes $\mathbf{u}(a,\cdot)$ on the budget set $\{x \text{ in } P, px = r\}$.

LEMMA 1. *Under assumptions* (A-1) *and* (A-2), *if* B *is a compact subset of* E, *then for all* p *in* S, *for all* $\bar{r} > \underline{r} > 0$, *and for all* $k > 0$, *there exists* $\varepsilon > 0$, *such that*

$$\textit{for all } (a,b) \textit{ in } B \times B \textit{ satisfying } \|b-a\| < \varepsilon,$$

$$\{\bar{r} \geqslant r' > r + k\|b-a\| \geqslant r > \underline{r}\} \quad \textit{implies} \quad \{\mathbf{u}(a, \xi(b,p,r')) > \mathbf{u}(a, \xi(a,p,r))\}.$$

Proof of Lemma 1. We first prove three claims.

Claim 1. The demand function $\xi(a,p,r)$ is locally lipschitzian with respect to a, and continuously differentiable with respect to r.

---

[4] Let us denote by $\mathscr{P}$ the set of all monotonic and continuous preference relations on P. Let a and b be two elements of $\mathscr{P}$ with their corresponding Kannaï reresentations $u(a,\cdot)$ and $u(b,\cdot)$. Define $d(a,b) = \sup_{x \text{ in } P} |u(a,x) - u(b,x)|/(1 + \|x\|)$. It is easy to check that d is a distance on $\mathscr{P}$ which induces a topology on $\mathscr{P}$ corresponding to our assumption (A-4). Let us recall that the topology of closed convergence on $\mathscr{P}$ introduced by Kannaï [1970] and Hildenbrand [1970] is metrizable with the distance δ defined as: $\delta(a,b) = \sup_{x \text{ in } P} |u(a,x) - u(b,x)|/(1 + \|x\|^2)$. It is therefore clear that the topology induced by the metric d is finer than the topology of closed convergence. However, it can be shown that the two topologies are equivalent on the set of homothetic preferences. Whether the set $\mathscr{P}$ endowed with the metric d is separable is an open question.

In fact, by Lagrange, $\xi(a,p,r)$ is the unique vector x which satisfies the following system of equations

$$D\mathbf{u}(a,x) - \lambda p = 0,$$

$$px - r = 0,$$

in the unknowns (x, λ). The differentiability with respect to r follows from the standard argument of Debreu [1972]. The fact that $\xi(a,p,r)$ is locally lipschitzian can be derived from a similar argument, using Clarke's generalization of the inverse function theorem for lipschitzian mappings (see Cornet and Laroque [1980] for details).

Claim 2. For all p in S and all $\bar{r} > \underline{r} > 0$, ther exists a strictly positive ρ such that:

$$\text{for all } a \text{ in } B, \qquad \text{if} \quad \bar{r} \geqslant r' \geqslant r \geqslant \underline{r}, \qquad \text{then}$$

$$\mathbf{u}(a, \xi(a,p,r')) \geqslant \mathbf{u}(a, \xi(a,p,r)) + \rho(r' - r).$$

This follows immediately from the fact that the function $\mathbf{u}(a, \xi(a,p,\cdot))$ is differentiable and that its derivative is continuous with respect to a and r and strictly positive on the compact set $B \times [\underline{r}, \bar{r}]$.

Claim 3. For all p in S, all $\bar{r} > \underline{r} > 0$ and all $h > 0$, there exists $\varepsilon > 0$ such that

$$\text{for all } (a,b) \text{ in } A \times A, \qquad \text{for all } r, \bar{r} > r > \underline{r} \quad \text{and} \quad \|b - a\| < \varepsilon \qquad \text{imply}$$

$$\mathbf{u}(a, \xi(a,p,r)) - \mathbf{u}(a, \xi(b,p,r)) \leqslant h\,\|b - a\|.$$

Consider the function

$$\begin{aligned} \psi(r,a,b) &= |\mathbf{u}(a, \xi(b,p,r) - \mathbf{u}(a,p,r))|/\|b - a\| && \text{for} \quad b \neq a \\ &= 0 && \text{for} \quad b = a \end{aligned}$$

This is a well-defined function for all $r > 0$ and (a,b) in $E \times E$.

To show that ψ is continuous, consider the second order Taylor development.

$$\begin{aligned} \mathbf{u}(a, \xi(b,p,r)) - \mathbf{u}(a, \xi(a,p,r)) &= D\mathbf{u}(a, \xi(a,p,r))[\xi(b,p,r) - \xi(a,p,r)] \\ &\quad + {}^{t}[\xi(b,p,r) - \xi(a,p,r)]\, D^2\mathbf{u}(a, \bar{\xi})[\xi(b,p,r) - \xi(a,p,r)], \end{aligned}$$

where $D^2\mathbf{u}(a,x)$ is the matrix of second order derivatives of $\mathbf{u}(a,x)$ with respect to x and $\bar{\xi}$ a convex combination of $\xi(b,p,r)$ and $\xi(a,p,r)$. The first term of the development is equal to zero since $D\mathbf{u}(a, \xi(a,p,r)) = \lambda p$ and $p\xi(a,p,r) = p\xi(b,p,r) = r$. Thus the lipschitzian character of $\xi(a,p,r)$ with respect to a (claim 1) implies that $\psi(r, a, b)$ tends to zero when a tends to b. Therefore ψ is continuous.

Now, for all r, a, b:

$$\psi(r, a, a) = 0,$$

$$\psi(r, a, b) \leqslant 0.$$

Since ψ is uniformly continuous on the compact set $[\underline{r}, \bar{r}] \times B \times B$, for all $h > 0$ there exists $\varepsilon > 0$ such that $\bar{r} \geqslant r \geqslant \underline{r}$ and $\|b - a\| < \varepsilon$ implies $|\psi(r, a, b) - \psi(r, a, a)| = -\psi(r, a, b) \leqslant h$. This completes the proof of claim 3.

Now take k, $\underline{r}$, $\bar{r}$, r, r' satisfying the conditions given in the statement of the lemma.

$$\mathbf{u}(a, \xi(b, p, r')) - \mathbf{u}(a, \xi(a, p, r)) = \mathbf{u}(a, \xi(b, p, r')) - \mathbf{u}(a, \xi(a, p, r')) \\ + \mathbf{u}(a, \xi(a, p, r')) - \mathbf{u}(a, \xi(a, p, r)).$$

By claim 2, the last difference is superior or equal to $\rho(r' - r) > \rho k \|b - a\|$. Thus it suffices to take h strictly inferior to ρk and the corresponding ε in claim 3 to obtain the desired result. Q.E.D.

LEMMA 2. *Under assumption* (A-4), *for all a in E, for all p in S and for all $\bar{u}$ in $\mathscr{R}$ there exist strictly positive η and $\bar{r}$ such that*

$$\{\|b - a\| < \eta \text{ and } r > \bar{r}\} \qquad \text{implies} \qquad \{\mathbf{u}(a, \xi(b, p, r)) > \bar{u}\}.$$

Proof of Lemma 2. First, since $\mathbf{u}(b, \cdot)$ is a Kannaï representation of the preference relation, $\mathbf{u}(b, \xi(b, p, r))$ is larger than the distance between the origin and the intersection of the line $\{\lambda e, \lambda > 0\}$ with the budget set. Thus:

$$\mathbf{u}(b, \xi(b, p, r)) \geqslant r/(pe).$$

Second, since $p \cdot \xi(b, p, r) = r$, each coordinate of $\xi(b, p, r)$ is inferior or equal to r divided by the smallest coordinate of p. Thus given p, there exists $K > 0$ such that for all b in E, for all r, $\|\xi(b, p, r)\| \leqslant Kr$.

Now, by assumption (A-4), taking $\varepsilon = 1/2(pe)\, K$, there exists $\eta > 0$ such that if $\|b - a\| < \eta$, then

$$\mathbf{u}(a, x) \geqslant \mathbf{u}(b, x) - (1 + \|x\|/2(pe)\, K).$$

Take $x = \xi(b, p, r)$ and apply the two inequalities given above:

$$\mathbf{u}(a, \xi(b, p, r)) \geqslant r/(pe) - 1/(2(pe)\, K) - r/2(pe).$$

If we choose $\bar{r} = 2(pe)\, \bar{u} + 1/K$, we obtain the desired result. Q.E.D.

We are now in a position to undertake the proofs of Proposition 1 and of the theorem. From now on, $\mathbf{x}$ is a given fair allocation and $\mathbf{r}$ the corresponding income distribution. The proofs use two lemmas. In the first lemma, we construct a subset A_1 of A where the property of fairness is satisfied everywhere.

LEMMA 3. *Under assumption* (A-1), *there exists a measurable subset* A_1 *of* A *and a price system* p *in* S, *such that*

(i) $\mu(A_1) = 1$,

(ii) $\mathbf{x}(a) = \xi(a, p, \mathbf{r}(a))$ *for all* a *in* A_1,

(iii) $\mathbf{u}(a, \mathbf{x}(b)) \leqslant \mathbf{u}(a, \mathbf{x}(a))$ *for all* (a, b) *in* $A_1 \times A_1$.

Furthermore there exists $\underline{r} > 0$ *such that*

(iv) $\mathbf{r}(a) \geqslant \underline{r}$ *for all* a *in* A_1.

Proof of Lemma 3. First, it is well known that there exists a price system p associated with every Pareto-optimal allocation such that $\mathbf{x}(a) = \xi(a, p, \mathbf{r}(a))$ except on a set B_1 of measure zero.

Second, let B_2 be the set of envious agents, i.e., the set of agents a in A such that $\mu(\{b \text{ in } A,\ \mathbf{u}(a, \mathbf{x}(b)) > \mathbf{u}(a, \mathbf{x}(a))\}) > 0$. The definition of an equitable allocation and Fubini's theorem imply that $\mu(B_2) = 0$.

Finally, let ν be the image measure defined on P by $\nu(\cdot) = \mu(\mathbf{x}^{-1}(\cdot))$, and let P_ν be the support of ν (see Parthasarathy [1967, p. 27, Theorem 2.1]). Let B_3 be the set of agents a such that $\mathbf{x}(a)$ does not belong to P_ν. By definition of P_ν, B_3 is of measure zero.

Define A_1 as the complement in A of B_1 union B_2 union B_3. Clearly (i) and (ii) of the lemma are satisfied.

We now prove that (iii) is satisfied. Suppose a contrario that there exists (a, b) in $A_1 \times A_1$ such that $u(a, \mathbf{x}(b)) > u(a, \mathbf{x}(a))$. Then the set $\{x \in P,\ u(a, x) > u(a, \mathbf{x}(a)\}$ which is open by continuity of $\mathbf{u}(a, \cdot)$, has a nonempty intersection with P_ν, since b does not belong to B_3. But it is a property of the support P_ν that any open set in P which has a non empty intersection with P_ν has a strictly positive measure. Thus a is envious, a contradiction with the fact that it does not belong to B_2.

Finally, we prove (iv). Take any element b in A_1 and let $\underline{r} = \mathrm{Min}_h\, p_h \mathbf{x}_h(b)$. We show that (iv) holds by reductio ad absurdum. Suppose that there exists a in A_1 with $\mathbf{r}(a) < \underline{r}$. The budget constraint implies:

$$p_h \mathbf{x}_h(a) \leqslant \mathbf{r}(a) < \underline{r} \leqslant p_h \mathbf{x}_h(b).$$

Thus $\mathbf{x}(a)$ has all its coordinates smaller than $\mathbf{x}(b)$, and by monotonicity, $\mathbf{u}(a, \mathbf{x}(a)) < \mathbf{u}(a, \mathbf{x}(b))$, a contradiction with (iii). Q.E.D.

The main step of the proof is the following lemma.

LEMMA 4. *Let a_0 be a point in the interior of A and assume that there exists a compact ball $B(a_0, \eta)$ of center a_0 and radius η contained in A and a constant $\bar{r}$ such that*

$$\mathbf{r}(a) \leqslant \bar{r} \qquad \textit{for all } a \textit{ in } A_1 \cap B(a_0, \eta).$$

Under assumptions (A-1) *and* (A-2), $\mathbf{r}$ *is equal to a constant function on* $A_1 \cap B(a_0, \eta)$.

Proof of Lemma 4. Since A is the support of μ and $\mu(A_1) = \mu(A) = 1$, A_1 is dense in A (in fact any open set which has a nonempty intersection with A must have a strictly positive measure).

Now, by Lemma 1, the function $\mathbf{r}$ restricted to $A_1 \cap B(a_0, \eta)$ is uniformly continuous. Therefore there exists a unique function $\mathbf{s}$ defined on $B(a_0, \eta)$ which is uniformly continuous and such that (see Schwartz [1970, Theorem 2, XI, 3, 1]):

$$\mathbf{s}(a) = \mathbf{r}(a) \qquad \text{on} \qquad A_1 \cap B(a_0, \eta).$$

Let $\mathbf{y}(a) = \xi(a, p, \mathbf{s}(a))$. By continuity, we have

$$\mathbf{u}(a, \mathbf{y}(b)) \leqslant \mathbf{u}(a, \mathbf{y}(a)) \qquad \text{for all } a \text{ and } b \text{ in } B(a_0, \eta).$$

We shall show that the function $\mathbf{s}$ is constant on $B(a_0, \eta)$. Suppose, a contrario, that there exists b in $B(a_0, \eta)$ such that $|\mathbf{s}(b) - \mathbf{s}(a_0)| > 0$. Let

$$k = |\mathbf{s}(b) - \mathbf{s}(a_0)|/2\, \|b - a_0\|,$$

and take the corresponding ε associated with k through Lemma 1. Let n be the integer part of $|\mathbf{s}(b) - \mathbf{s}(a_0)|/\varepsilon$, and define $a_j = a_0 + j\varepsilon(b - a_0)/\|b - a_0\|$, $j = 1,..., n$, $a_{n+1} = b$. Now, $\mathbf{u}(a_j, \mathbf{y}(a_{j+1})) \leqslant \mathbf{u}(a_j, \mathbf{y}(a_j))$ implies, by Lemma 1, $|\mathbf{s}(a_{j+1}) - \mathbf{s}(a_j)| \leqslant k\, \|a_{j+1} - a_j\|$.

Therefore: $|\mathbf{s}(a_0 - \mathbf{s}(b)| \leqslant k\, \|a_0 - b\|$, a contradiction with the definition of k. Q.E.D.

Proofs of Proposition 1 *and of the theorem.* We give the proof of the theorem. Claim 1 deals with boundedness problems[5] which are resolved by

---

[5] If the assumption (A-4) is fulfilled and if A is compact, one can show that the income distribution corresponding to a fair allocation is essentially bounded. Then, one can associate with every fair allocation $\mathbf{x}$ another feasible allocation $\mathbf{y}$ such that:

$$\begin{aligned} &\mathbf{y}(a) = \mathbf{x}(a), && \mu\text{.a.e. in } A,\\ &\mathbf{y}(a) = \xi(a, p, p \cdot \mathbf{y}(a)), && \text{for } \textit{all } a \text{ in } A,\\ &\mathbf{u}(a, \mathbf{y}(a)) \geqslant \mathbf{u}(a, \mathbf{y}(b)), && \text{for } \textit{all } (a, b) \text{ in } A \times A. \end{aligned}$$

In other words, the properties which are satisfied almost everywhere for xhold everywhere for $\mathbf{y}$. In this case a shorter proof of the theorem could be given.

assumption in Proposition 1. Step 2 can be applied directly as a self contained proof of Proposition 1.

Claim 1. Under assumption (A-4), for all a in A, there exists a compact ball $B(a, \eta)$ and a constant $\bar{r}$ such that

$$\mathbf{r}(b) \leqslant \bar{r} \qquad \text{for all } b \text{ in } A_1 \cap B(a, \eta).$$

First note that the claim is valid for a in A_1. It suffices to apply Lemma 2 with $\bar{u} = \mathbf{u}(a, \xi(a, p, \mathbf{r}(a)))$, together with the property of nonenvy on A_1.

Take now any a in A and $B(a, \eta)$ an arbitrary compact ball. For all b in $A_1 \cap B(a, \eta)$ there exists an open ball where $\mathbf{r}$ is bounded on A_1. Since A_1 is dense in A, all these balls form a covering of $B(a, \eta) \cap A$ and by compacity, there exists a finite covering of $B(a, \eta) \cap A$ by such balls. Thus $\mathbf{r}(b)$ is bounded on $A_1 \cap B(a, \eta)$. This completes the proof of claim 1.

Step 2. Define a function $\mathbf{s}$ on the interior of A by the following device. For every a in the interior of A apply Lemma 4 by choosing an appropriate value of η (which is always possible by assumption in Proposition 1, or by claim 1 above in the theorem), and let $\mathbf{s}(a)$ be the constant value taken by $\mathbf{r}$ on $A_1 \cap B(a, \eta)$. Clearly $\mathbf{s}$ is continuous, and equal to $\mathbf{r}$ on the intersection of A_1 and of the interior of A.

Let $\bar{s}$ be a value taken by $\mathbf{s}$. Now the set $\{a$ in the interior of A, $\mathbf{s}(a) = \bar{s}\}$ is closed in the interior of A by continuity of $\mathbf{s}$. It is open by Lemma 4. Since by assumption (A-3), the interior of A is connected, this set is equal to the interior of A. Thus $\mathbf{r}$ is constant on the intersection of A_1 and of the interior of A.

Finally note that by claim 1 above and Lemma 1, $\mathbf{r}$ is locally bounded and, thus, continuous on A_1. Since A is the closure of its interior, $\mathbf{r}$ is constant on A_1. Q.E.D.

REFERENCES

CORNET, B. AND LAROQUE, G. [1980], "Lipschitz properties of constrained demand functions and constrained maximizers," ENSAE, Discussion Paper.

DEBREU, G. [1969], Neighboring economic agents in "La Décision," pp. 85-90, Éditions du C.N.R.S., Paris.

DEBREU, G. [1972], Smooth preferences, *Econometrica* **40**, 603–615.

HAMMOND, P. J. [1979], Straightforward individual incentive compatibility in large economies, *Rev. Econ. Studies* **46**, 263–282.

HILDENBRAND, W. [1970], On economies with many agents, *J. Econ. Theory* **2**, 161–188.

KANNAI, Y. [1970], Continuity properties of the core of a market, *Econometrica* **38**, 791–815.

PARTHASARATHY, K. R. [1967], "Probability Measures on Metric Spaces," Academic Press, New York.

SCHMEIDLER, D. AND VIND, K. [1972], Fair net trades, *Econometrica* **40**, 637–642.
SCHWARTZ, L. [1970], "Analyse, Topologie Générale et Analyse Fonctionelle," Hermann, Paris.
VARIAN, H. R. [1976], Two problems in the theory of fairness, *J. Public Econ.* **5**, 249–260.

JOURNAL OF ECONOMIC THEORY 25, 283–286 (1981)

Notes, Comments, and Letters to the Editor

Super-modularity: Applications to Convex Games and to the Greedy Algorithm for LP*

TATSURO ICHIISHI

Departments of Economics and Mathematics,
The University of Iowa, Iowa City, Iowa 52242

Received October 17, 1978; revised August 15, 1980

L. S. Shapley (*Internat. J. Game Theory* 1 (1971), 11–26) showed that if a game is convex, then all the marginal worth vectors are members of the core. J. Edmonds (*in* "Combinatorial Structures and Their Applications" (R. Guy *et al.*, Eds.), pp. 69–87, Gordon & Breach, New York, 1970) considered a class of linear programming problems for which the greedy algorithm works. The present paper unifies these studies, and establishes the converse of each theorem.

1

Shapley [3] showed that if a game is convex, then all the marginal worth vectors are members of the core [in fact, they are precisely the vertices of the core]. The present author established the converse of the Shapley theorem in the original version of this note, and subsequently was informed by an anonymous referee that the study of convex games is closely related to Edmonds' study of the greedy algorithm for linear programming in [1]. The purpose of the present version is to unify all these studies. Related work can be found in Weber [4].

2

Let N be a finite set of integers $\{1, 2, \dots, n\}$, and $\mathcal{N}$ the power set of N. A real-valued function on $\mathcal{N}$, $v: \mathcal{N} \to \mathbf{R}$, is called *super-modular*, if $v(T \cup S) +$

---

* A revised version of: T. Ichiishi, The core and the marginal worth vectors, mimeo, January 1978. The author would like to express his deep gratitude to an anonymous referee for pointing out the need for a revision. The author's research is supported in part by U.S. National Science Foundation Grant SES 7806123 (formerly SOC 78-06123).

0022-0531/81/050283-04$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

$v(T\cap S)\geqslant v(T)+v(S)$, for any $T, S\in\mathcal{N}$. Given a permutation σ on N, define $P^\sigma_j\equiv\{i\in N\mid\sigma(i)<\sigma(j)\}$; the set of members in N which precede j with respect to the order σ. Denote by G_n the symmetric group, i.e., the set of all permutations on N. Given any $\sigma\in G_n$ and any function $v\colon\mathcal{N}\to\mathbf{R}$, define $a^\sigma(v)\in\mathbf{R}^n$ by: $a^\sigma_j(v)\equiv v(P^\sigma_j\cup\{j\})-v(P^\sigma_j)$, for every $j\in N$.

THEOREM. *Let $v\colon\mathcal{N}\to\mathbf{R}$ be such that $v(\phi)=0$. Then, the following four conditions are equivalent:*

(i) *The function v is super-modular;*

(ii) *For any $A, B, C\in\mathcal{N}$ such that $A\subset B\subset N\backslash C$,*

$$v(B\cup C)-v(B)\geqslant v(A\cup C)-v(A);$$

(iii) *For any $A, B\in\mathcal{N}$ and any $j\in N$ such that $A\subset B\subset N\backslash\{j\}$,*

$$v(B\cup\{j\})-v(B)\geqslant v(A\cup\{j\})-v(A);$$

(iv) *For any $\sigma\in G_n$ and for any $S\in\mathcal{N}$,*

$$\sum_{j\in S}a^\sigma_j(v)\geqslant v(S).$$

Proof. (i) ⇒ (ii) ⇒ (iii): Trivial.

(iii) ⇒ (iv): Put $S\equiv\{j_1,j_2,\dots,j_s\}$, so that $\sigma(j_1)<\sigma(j_2)<\cdots<\sigma(j_s)$. Apply (iii) for $[A=\{j_1,\dots,j_{i-1}\}, B=P^\sigma_{j_i}, j=j_i]$, or $v(P^\sigma_{j_i}\cup\{j_i\})-v(P^\sigma_{j_i})\geqslant v(\{j_1,\dots,j_i\})-v(\{j_1,\dots,j_{i-1}\})$. (j_i).

Summing up the inequalities (j_1)–(j_s), one obtains the required results.

(iv) ⇒ (i): Given any $T,S\in\mathcal{N}$, put $T\cap S\equiv\{j_1,\dots,j_r\}$; $T\backslash S\equiv\{j_{r+1},\dots,j_s\}$; and $S\backslash T\equiv\{j_{s+1},\dots,j_t\}$, where $0\leqslant r\leqslant s\leqslant t\leqslant n$. One may, therefore, put $N\backslash(T\cup S)\equiv\{j_{t+1},\dots,j_n\}$. Now, consider a permutation $\sigma\in G_n$ defined by $\sigma\colon j_i\mapsto i$. Then, $P^\sigma_{j_i}=\{j_1,\dots,j_{i-1}\}$ for each $i\in N$, or

$$\begin{aligned}&\sum_{j\in S}[v(P^\sigma_j\cup\{j\})-v(P^\sigma_j)]\\&\quad=\sum_{i=s+1}^{t}[v(T\cup\{j_{s+1},\dots,j_i\})-v(T\cup\{j_{s+1},\dots,j_{i-1}\})]\\&\qquad+\sum_{i=1}^{r}[v(\{j_1,\dots,j_i\})-v(\{j_1,\dots,j_{i-1}\})]\\&\quad=[v(T\cup S)-v(T)]+[v(T\cap S)-v(\phi)].\end{aligned}$$

By (iv), this is greater than or equal to $v(S)$. Q.E.D.

Remark. As will be pointed out in Sections 3 and 4, the only new result here is the statement [(iv) $\Rightarrow$ (i)].

3

A *game* (or more precisely, a *cooperative game in characteristic function form with sidepayments*) is a function $v: \mathcal{N} \to \mathbf{R}$, such that $v(\phi) = 0$. The set N is interpreted as the set of players, and $\mathcal{N}$ as the family of coalitions. The *core* of a game v is the set of outcomes which are feasible and stable; i.e., the set $\{x \in \mathbf{R}^n \mid \sum_{j \in N} x_j \leqslant v(N);\ \sum_{j \in S} x_j \geqslant v(S)$ for every $S \in \mathcal{N}\}$. The core is a convex polyhedron. Given a permutation σ on N, the *marginal worth vector* of a game v is simply the vector $a^\sigma(v)$ as defined in Section 2. Note that $\# G_n = n!$. The *Shapley value* is a map, $v \mapsto \sum_{\sigma \in G_n} (1/n!)\, a^\sigma(v)$. [See Shapley [2].] A game v is called *convex*, if the function v is super-modular. A game v is said to satisfy *increasing returns with respect to the coalition size*, if for any $A, B \in \mathcal{N}$ and any $j \in N$ such that $A \subset B \subset N \backslash \{j\}$, $v(B \cup \{j\}) - v(B) \geqslant v(A \cup \{j\}) - v(A)$. It is easy to verify that, for a marginal worth vector $a^\sigma(v)$, $\sum_{j \in N} a_j^\sigma(v) = v(N)$. The statements [(1.i) $\Leftrightarrow$ (1.ii) $\Rightarrow$ (1.iii)] of the following Corollary 1 are due to Shapley [3, p. 13, and Theorem 4]:

COROLLARY 1. *Let $v: \mathcal{N} \to \mathbf{R}$ be a game. Then, the following three conditions are equivalent*:

(1.i) *The game v is convex*;

(1.ii) *The game v satisfies increasing returns with respect to the coalition size*;

(1.iii) *All the marginal worth vectors of v, $a^\sigma(v)$, $\sigma \in G_n$ are members of the core of v.*

Remark. Shapley [3, Theorems 3 and 5] and Edmonds [1, Theorem (22)] independently established that if the condition (1.i) holds, then the set $\{a^\sigma(v) \mid \sigma \in G_n\}$ is precisely the set of vertices of the core. Weber [4] has shown that the convex hull of the set $\{a^\sigma(v) \mid \sigma \in G_n\}$ contains the core of *any* game v.

4

Given any point $c \in \mathbf{R}^n$, and any function $w: \mathcal{N} \to \mathbf{R}$, consider the following linear programming problem: (P) Maximize $c \cdot x$, subject to $\sum_{j \in S} x_j \leqslant w(S)$ for every $S \in \mathcal{N}$, $x \geqslant \mathbf{0}$. Choose $\sigma \in G_n$ so that $c_{\sigma(1)} \geqslant$

$c_{\sigma(2)} \geqslant \cdots \geqslant c_{\sigma(k)} > 0 \geqslant c_{\sigma(k+1)} \geqslant \cdots$. For each integer i, $1 \leqslant i \leqslant k$, put $A_i \equiv \{\sigma(1), \sigma(2), \ldots, \sigma(i)\}$. To use the notation in Section 2, $A_i = P^{\sigma^{-1}}_{\sigma(i)} \cup \{\sigma(i)\}$. One says that the *greedy algorithm works* for (P), if the vector $x^\sigma \in \mathbf{R}^n$, defined by: $x^\sigma_{\sigma(1)} \equiv w(A_1)$; $x^\sigma_{\sigma(i)} \equiv w(A_i) - w(A_{i-1})$ for $2 \leqslant i \leqslant k$; $x^\sigma_{\sigma(i)} \equiv 0$ for $k < i \leqslant n$, is an optimal solution of (P). The statement [(2.iii) ⇒ (2.i)] of the following Corollary 2 is due to Edmonds [1, Theorem (19)], and one easy proof would be to use the Theorem of Section 2 and the duality theorem of LP. Clearly, [(2.i) ⇒ (2.ii)].

COROLLARY 2. *Let* $w: \mathcal{N} \to \mathbf{R}$ *be such that* $w(\phi) = 0$. *Then, the following three conditions are equivalent*:

(2.i) *The greedy algorithm works for any* $c \in \mathbf{R}^n$;

(2.ii) *The greedy algorithm works for* $c = \mathbf{1}$, *the vector each of whose components is* 1;

(2.iii) *The function* w *is non-negative-valued, non-decreasing* (*i.e.*, $S \subset T \Rightarrow w(S) \leqslant w(T)$), *and sub-modular* (*i.e.*, $-w$ *is super-modular*).

Proof. (2.ii) ⇒ (2.iii): By choosing suitable permutations, one can show from non-negativity of x^σ that w is non-negative-valued and non-decreasing. Now, take any $\sigma \in G_n$. Since $x^\sigma_j = w(P^{\sigma^{-1}}_j \cup \{j\}) - w(P^{\sigma^{-1}}_j)$, the feasibility of x^σ implies $\sum_{j \in S} [w(P^{\sigma^{-1}}_j \cup \{j\}) - w(P^{\sigma^{-1}}_j)] \leqslant w(S)$ for every $S \in \mathcal{N}$. Apply the Theorem of Section 2, to get super-modularity of $-w$. Q.E.D.

REFERENCES

1. J. EDMONDS, Submodular functions, matroids, and certain polyhedra, *in* "Combinatorial Structures and their Applications" (R. Guy *et al.*, Eds.), pp. 69–87, Gordon & Breach, New York, 1970.
2. L. S. SHAPLEY, A value for n-person games, *in* "Contributions to the Theory of Games, II" (H. W. Kuhn, A. W. Tucker, Eds.), pp. 307–317, Princeton Univ. Press, Princeton, N. J., 1953.
3. L. S. SHAPLEY, Cores of convex games, *Internat. J. Game Theory* 1 (1971), 11–26.
4. R. J. WEBER, Probabilistic values for games, mineo, Yale University, 1978.

JOURNAL OF ECONOMIC THEORY 25, 287–301 (1981)

Giffen Inputs and the Theory of Multiple Production

JOSEPH P. HUGHES

Department of Economics, Rutgers College CN5055,
New Brunswick, New Jersey 08903

Received January 17, 1979; revised May 27, 1980

The effect on the single-product firm's profit-maximizing demand for an input of a change in its price can be broken down into a substitution effect, which is always negative, and an output effect, which always reinforces the substitution effect, even when the input is inferior ([4, 7]). Thus, the analogue to the Giffen good in consumption theory does not exist in the case of the single-product firm engaged in unconstrained profit maximization. Sakai [7] has extended this conclusion to the case of joint production to show that even when an input is employed to produce several products, its total employment is still inversely related to its price. His conclusion is valid, not just for joint production, but for all instances of unconstrained, profit-maximizing multiple production, and not just for joint inputs, but for nonjoint inputs as well.

However, as this paper shows, when the aggregate demand for a nonjoint input is broken down into the specific demands attributable to each product, Sakai's conclusion is not generally valid for the specific demands. The effect of a change in the input's price on the specific demand can be decomposed into a substitution and an output effect, but neither has a definite a priori sign. Thus, the Giffen phenomenon appears in the theory of production in an even more startling fashion than in the theory of consumption: trouble may arise not only from the output effect, but also from the substitution effect.

The concept of an input's specific demand and the possibility of a Giffen effect turn out to be of considerable importance for public policy. The discussion of this point is postponed, though, for Section VII. Section I reviews Sakai's results while Sections II–V examine the properties of the output-specific demands for inputs and compare them to Sakai's findings for the aggregate demands. A numerical example of a positive output-specific substitution effect is provided in Section VI.

0022-0531/81/050287-15$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

1. PRELIMINARY DEFINITIONS AND RELATIONSHIPS

Consider a firm producing m products at the rates $\bar{y} = (y_1, y_2, \ldots, y_m)$, selling them in competitive markets at the prices $\bar{p} = (p_1, p_2, \ldots, p_m)$, employing n inputs at the rates $\bar{x} = (x_1, x_2, \ldots, x_n)$, and purchasing them at the competitive prices $\bar{w} = (w_1, w_2, \ldots, w_n)$. If the jth input is a joint or "public" input, then its rate of employment in producing the ith product, x_{ij}, is the same as that in producing all others, $x_j = x_{ij}$ $(i = 1, \ldots, m)$. However, if the jth input is not joint, then $x_j = \sum_{i=1}^{m} x_{ij}$.

In the absence of any type of interdependency among the production processes or the demands for the products, the cost and profit functions could be expressed as separable functions of rates and of prices of outputs respectively:

$$\text{(i)} \quad C = \sum_{i=1}^{m} C_i(y_i, \bar{w}) \qquad \text{and} \qquad \text{(ii)} \quad \pi = \sum_{i=1}^{m} \pi_i(p_i, \bar{w}). \tag{1}$$

It is well known that, from their definitions, C_i is concave in $\bar{w}$, and π_i is convex in $(p_i, \bar{w})$.[1] Where there are interdependencies in production, neither function can be written separably:

$$\text{(i)} \quad C = C(\bar{y}, \bar{w}) \qquad \text{and} \qquad \text{(ii)} \quad \pi = \pi(\bar{p}, \bar{w}). \tag{2}$$

When the linkage is confined to demand and the firm sells in monopolistic markets, only the cost function can be written separably. As before, C is concave in $\bar{w}$, π is convex in $(\bar{p}, \bar{w})$, and both functions are assumed to be twice continuously differentiable.

There are at least four possible sources of interdependencies.

(i) There may be an input whose supply to the firm is fixed, but whose allocation among the various production processes is variable [5].

(ii) There may be an input whose use in one process does not diminish its use in another. In this sense it is a joint or "public" input.

(iii) The separate production of each product may entail external effects, which are internalized by the structure of multiple production.

(iv) There may be interdependent demands for the firm's outputs which, in the case of monopoly, explicitly tie together the decisions regarding rates of outputs. Unless otherwise stated, it is assumed that at least one of (i)–(iii) hold so that neither objective function is separable.

By Hotelling's lemma, the profit-maximizing demand for an input is found to be $-\partial\pi/\partial w_j = x_j(\bar{p}, \bar{w})$ and the supply of output, to be $\partial\pi/\partial p_k = y_k(\bar{p}, \bar{w})$.

---

[1] See, for example, [2; 7; 8, pp. 28–30]. Moreover, at the competitive firm's equilibrium, C_i must also be convex in y_i so that marginal cost is increasing in y_i.

By Shepard's lemma, the constant-output demand is $\partial C/\partial w_j = u_j(\bar{y}, \bar{w})$. Proofs of these lemmas are cited by Sakai [7, pp. 259, 261] and will not be duplicated here. The convexity of profit precludes the possibility of a Giffen effect in the aggregate demands:

$$\partial^2\pi/\partial w_j^2 = \partial x_j/\partial w_j \leqq 0. \tag{3}$$

Decomposing the total effect, Sakai proves that

$$\partial x_j(\bar{p}, \bar{w})/\partial w_j = \partial u_j(\bar{y}, \bar{w})/\partial w_j + \sum_{k=1}^{m} (\partial u_j(\bar{y}, \bar{w})/\partial y_k)(\partial y_k(\bar{p}, \bar{w})/\partial w_j).^2 \tag{4}$$

The aggregate substitution effect is nonpositive since, by the concavity of C,

$$\partial u_j/\partial w_j = \partial^2 C/\partial w_j^2 \leqq 0. \tag{5}$$

The aggregate output effect is also nonpositive. Observe that

$$\partial y_k/\partial w_j = \partial^2\pi/\partial w_j\partial p_k = \partial^2\pi/\partial p_k\partial w_j = -\partial x_j/\partial p_k \tag{6}$$

and

$$\partial x_j/\partial p_k = \sum_{i=1}^{m} (\partial u_j/\partial y_i)(\partial y_i/\partial p_k) \tag{7}$$

so that substituting (6) and (7) into the expression for the output effect in (4) yields

$$-\sum_{k=1}^{m}\sum_{i=1}^{m} (\partial u_j/\partial y_k)(\partial y_i/\partial p_k)(\partial u_j/\partial y_i), \tag{8}$$

which, due to the convexity of π and the consequent positive semidefiniteness of the matrix $[\partial y_i/\partial p_k]$, is nonpositive [7, p. 267].

II. Defining an Inferior Input

The theory of single (or separable multiple) production offers two pairs of interchangeable definitions of an inferior input:

$$\partial u_{kj}/\partial y_k = \partial u_j/\partial y_k = \partial(\partial C/\partial y_k)/\partial w_j < 0 \tag{9}$$

and

$$-\partial x_{kj}/\partial p_k = -\partial x_j/\partial p_k = \partial y_k/\partial w_j > 0, \tag{10}$$

---

[2] See Lemma 1, [7, pp. 262–263].

which follow from the symmetry of the second order cross-partials of the separable functions in (1). Following the procecure in (6) and (7), $\partial y_k/\partial w_j = -(\partial u_j/\partial y_k)(\partial y_k/\partial p_k)$ so that, as was shown by Bear [1], a rise in the price of an input which is inferior in the production of output k (i.e., $\partial u_j/\partial y_k < 0$) will cause marginal cost to decrease and, consequently, the profit-maximizing rate of output to increase. In likewise, a rise in the price of output will increase production and decrease the input's rate of employment.

When (7) is substituted into (6), it can be seen that, in the case of nonseparable multiple production, the response of the kth output to a change in the jth input's price depends, not only on the input's normality or inferiority in the production of k (i.e., on the sign of $\partial u_j/\partial y_k$), but also on its relationship to all the other products. Although the response of the marginal cost of y_k to a change in the input's price follows the pattern of separable production,

$$\partial u_j/\partial y_k = \partial(\partial C/\partial y_k)/\partial w_j, \tag{11}$$

this relationship is not sufficient to predict the response of y_k. The relation in (11) holds at given rates of output, but as the firm adjusts those rates, additional changes in the marginal costs are induced by the linkages in production among the outputs. This point was neglected by Sakai and turns out to be of some importance. When the interchangeable definitions of normal-inferior appearing in the theory of single production, based on the signs of either (9) or (10), are extended to the case of nonseparable multiple production, they are no longer found to be equivalent.

III. THE OUTPUT-SPECIFIC DEMAND FOR INPUTS

Having spelled out the results of Eqs. (3)–(8) above, Sakai concludes, "The possibility of the Giffen case is thus excluded in production theory. This is a remarkable contrast to consumption theory, in which the substitution and income effects may go in opposite directions" [7, p. 267]. Had he considered the aggregate demand for a nonjoint input broken into its output-specific demands, he would have discovered the inaccuracy of that statement. From (3) it can be noted that

$$\sum_{i=1}^{m} \partial x_{ij}/\partial w_j = \partial x_j/\partial w_j \leqq 0 \tag{12}$$

but not necessarily that $\partial x_{ij}/\partial w_j \leqq 0$. Consider the decomposition

$$\partial x_{ij}/\partial w_j = \partial u_{ij}/\partial w_j + \sum_{k=1}^{m} (\partial u_{ij}/\partial y_k)(\partial y_k/\partial w_j). \tag{13}$$

From (5),

$$\sum_{i=1}^{m} \partial u_{ij}/\partial w_j = \partial u_j/\partial w_j \leqq 0 \tag{14}$$

but not necessarily $\partial u_{ij}/\partial w_j \leqq 0$. Thus, the specific substitution efect may be positive. Moreover, although it can be concluded from (8) that $\sum_{i=1}^{m} \sum_{k=1}^{m} (\partial u_{ij}/\partial y_k)(\partial y_k/\partial w_j) \leqq 0$, nothing can be said a priori concerning the sign of the specific output effect in (13). Consequently, neither the specific substitution nor the specific output effect is necessarily well behaved. Therefore, Giffen effects cannot be ruled out in the production of individual products.

One additional bit of unruliness must be noted. Since in the case of a nonjoint input, $\partial u_{kj}/\partial y_k \neq \partial u_j/\partial y_k$ and $\partial x_{kj}/\partial p_k \neq \partial x_j/\partial p_k$, none of the definitions of "normal-inferior" applied at this level of aggregation, $\partial u_{kj}/\partial y_k$, $\partial(\partial C/\partial y_k)/\partial w_j$, $\partial y_k/\partial w_j$, and $\partial x_{kj}/\partial p_k$, turn out to be necessarily equivalent.

Somewhat more order is restored if the source of interdependence resides only in demand. Having made production separable, the cost function in (1-i) and the profit function in (2-ii) hold. Thus, the specific substitution effects have nonpositive signs, leaving only the specific output effects in doubt. In addition, since $\partial u_{ij}/\partial y_i = \partial u_j/\partial y_i = \partial(\partial C/\partial y_i)/\partial w_j$, the definitions of "normal-inferior" based on cost become equivalent.

The possibility of positive substitution effects within particular production processes is not without intuitive appeal. Consider a firm producing two outputs with three or more variable inputs. Facing a rise in w_1, it would like to reduce u_{11} and u_{21} for any given rates of outputs. However, if there is another input, say the third, which is fixed in its total supply to the firm but can be switched from the production of one product to the other, the firm's response may be altered by the competition between the product lines for the rationed third input. Suppose that, at the given rates of output and in the absence of rationing, u_1 and u_3 are competitive in the production of both products, that is to say $\partial \hat{u}_{13}/\partial w_1 > 0$ and $\partial \hat{u}_{23}/\partial w_1 > 0$ (where the caret indicates unrationed demand). Thus, increasing the employment of the third input in either line allows a saving of the now-more-expensive first input. However, the third input can be increased in only one of the lines and that increase must be offset by a decrease of the same amount in the other. Suppose an increase in the employment of the rationed input in the second line permits a greater saving of the first input than an equal increase of the rationed input in the first line. Then the rationed input will be transferred from the production of the first to the second output; and while the increase in u_{23} will facilitate a reduction in the competitive u_{21}, it will necessitate a decrease in the rationed u_{13}. Since u_{11} and u_{13} are net substitutes in the absence of rationing, the decrease in u_{13} will create a tendency to increase

u_{11} which may overwhelm the usual tendency in the absence of rationing to reduce u_{11}. Thus, the competitive u_{11} may have to be raised to maintain the same rates of outputs. Of course, the saving of u_{21} must at least offset any added employment of u_{11}. Indeed, the motivation for the transfer of the rationed input from the first to the second line suggests such a saving. Thus, the aggregate substitution effect is nonpositive even though one of its components may be positive.[3] Similar lines of reasoning may be applied to the cases of joint inputs and internalized "externalities," except that the inputs in question must be competitive in the production of one product and complementary in the other. These assertions are proved in the following two sections and a numerical example is provided in the third.

The possibility of unruly output effects within particular production processes is more easily explained. Continuing to draw on the example above, assume that when outputs vary and there are no interdependencies, x_{11} and x_{21} are found to be normal in their respective processes. Then when their price rises, the marginal costs of both products increase, and the rates of outputs and employment, x_{11} and x_{21}, decline. On the other hand, if there are links on the side of production, such as rationing, the a priori relation between $\partial u_{ij}/\partial y_i$ and $\partial(\partial C/\partial y_i)/\partial w_j$ vanishes. For any given vector of outputs, the rise in w_1 causes both marginal costs to rise, but as either output is varied, the marginal cost of the other is again displaced. Because the adjustment in rates of output can lower as well as further raise marginal costs, it is not possible to place an a priori restriction on the sign of the specific output effects. A clearer example may be found in the case of separable production but interdependent demand. As noted above, a rise in w_1 increases the marginal cost of both products, but as either output is varied, the marginal revenue of the other is affected. Thus, a reduction in one output may increase the marginal revenue of the other such that it offsets the increased marginal cost and both its rate of production and the employment of the now more expensive input increase.

IV. OUTPUT-SPECIFIC SUBSTITUTION EFFECTS UNDER RATIONING

Following the example in the previous section, consider two well-behaved, *independent* production processes $y_i = f_i(x_{i1}, x_{i2}, x_{i3})$, $i = 1, 2$, consisting of three variable inputs. Minimizing cost given the vectors of competitive prices $\bar{w}$ and of outputs $\bar{y}$ yields the constant-output factor demands

$$\hat{u}_{ij} = \hat{u}_{ij}(\bar{w}, y_i) \tag{15}$$

---

[3] The impact of rationing on the multi-product firm is examined in some detail by Hughes [3].

and the cost function

$$C(\bar{w}, \bar{y}) = \sum_i C_i(\bar{w}, y_i) \tag{16}$$

in which each C_i is concave in $\bar{w}$ so that $\partial \hat{u}_{i1}/\partial w_1 \leqq 0$.

However, if the third input is rationed to the firm, the minimization of cost must include the constraint $\sum_i x_{i3} = x_3$ so that the cost function is no longer separable in y:

$$C(\bar{w}, \bar{y}, x_3) = \sum_i \sum_j w_j u_{ij}(w_1, w_2, \bar{y}, x_3). \tag{17}$$

The caret on u_{ij} is omitted to indicate rationed demand, and w_3 is dropped from the arguments of the demand function since it is assumed the constraint is binding at the value x_3. As was pointed out above, $\sum_i \partial u_{i1}/\partial w_1 \leqq 0$ but $\partial u_{i1}/\partial w_1 \gtreqless 0$. To demonstrate the condition under which this anomaly may arise, Robert Pollak's concept of a conditional demand function [6] is adapted to this context.

Let us assume that at the given values $(\bar{w}, \bar{y})$, the constrained value of x_3 is optimal:

$$x_3 = \hat{u}_3(\bar{w}, \bar{y}). \tag{18}$$

Then the unrationed demands can be expressed as $\hat{u}_{ij} = u_{ij}(w_1, w_2, \bar{y}, \hat{u}_3(\bar{w}, \bar{y}))$; consequently, for $j = 1$

$$\partial \hat{u}_{i1}/\partial w_1 = \partial u_{i1}/\partial w_1 + (\partial u_{i1}/\partial \hat{u}_3)(\partial \hat{u}_3/\partial w_1) \leqq 0 \tag{19}$$

by the concavity of C_i in $\bar{w}$. However, for the variation subject to rationing,

$$\partial u_{i1}/\partial w_1 = \partial \hat{u}_{i1}/\partial w_1 - (\partial u_{i1}/\partial \hat{u}_3)(\partial \hat{u}_3/\partial w_1) \gtreqless 0, \tag{20}$$

whose sign is ambiguous when the second term of the right member is negative.

To interpret the condition under which a positive specific substitution effect could arise, note that $\partial \hat{u}_{i1}/\partial w_3 = (\partial u_{i1}/\partial \hat{u}_3)(\partial \hat{u}_3/\partial w_3)$ so that solving for $\partial u_{i1}/\partial \hat{u}_3$ and substituting it and the identity $\partial \hat{u}_3/\partial w_1 \equiv \partial(\hat{u}_{13} + \hat{u}_{23})/\partial w_1$ into $\partial u_{i1}/\partial w_1$ results in the expression

$$\partial u_{i1}/\partial w_1 = \partial \hat{u}_{i1}/\partial w_1 - \partial \hat{u}_{i1}/\partial w_3(\partial \hat{u}_{13}/\partial w_1 + \partial \hat{u}_{23}/\partial w_1)/\partial \hat{u}_3/\partial w_3. \tag{21}$$

Letting $i = 1$, computing a common denominator, and substituting $\partial \hat{u}_3/\partial w_3 = \partial(\hat{u}_{13} + \hat{u}_{23})/\partial w_3$,

$$\begin{aligned}\partial u_{11}/\partial w_1 = \{[&(\partial \hat{u}_{11}/\partial w_1)(\partial \hat{u}_{13}/\partial w_3) \\ &- (\partial \hat{u}_{11}/\partial w_3)(\partial \hat{u}_{13}/\partial w_1)] + (\partial \hat{u}_{11}/\partial w_1)(\partial \hat{u}_{23}/\partial w_3) \\ &- (\partial \hat{u}_{11}/\partial w_3)(\partial \hat{u}_{23}/\partial w_1)\}/\partial \hat{u}_3/\partial w_3.\end{aligned} \tag{22}$$

The expression contained in the square brackets is a second-order principal minor of the Hessian matrix of the separable cost function C, and hence, is positive. The second term, a product of two negative elements on the main diagonal, is also positive. These terms, divided by the negative $\partial\hat{u}_3/\partial w_3$, would result in a negative specific substitution effect were it not for the last term in the braces.

Because C_1 is twice continuously differentiable, $\partial\hat{u}_{11}/\partial w_3 = \partial\hat{u}_{13}/\partial w_1$ so that the last term can be written as

$$(\partial\hat{u}_{11}/\partial w_3)(\partial\hat{u}_{23}/\partial w_1) = (\partial\hat{u}_{13}/\partial w_1)(\partial\hat{u}_{23}/w_1), \tag{23}$$

which is negative when the first and third inputs are competitive in one production line and complementary in the other. Thus, the anomalous substitution effect may occur only when the first and third inputs are either complementary in both lines or competitive in both.

In the previous section, the plausibility argument for a positive substitution efect in the first product line ($\partial u_{11}/\partial w_1 > 0$) relied on the suppositions that the first and third inputs are substitutes in both lines, that the increase in the price of the first can, under certain circumstances, cause the third to be transferred from the first to the second line, and that the first line's reduced ration may induce an increased employment of the now-more-expensive first input in that line.

In this section, the first supposition of the plausibility argument has been demonstrated correct: a necessary but not sufficient condition for the positive substitution effect is that the first and third inputs be either substitutes in both lines or complements. The next step in the argument supposed a transfer of the rationed input from the first to the second line ($\partial u_{23}/\partial w_1 > 0$). To determine the necessary and sufficient condition for such a shift, the procedure used to derive (21) is followed to arrive at

$$\partial u_{23}/\partial w_1 = \partial\hat{u}_{23}/\partial w_1 - (\partial\hat{u}_{23}/\partial w_3)(\partial\hat{u}_3/\partial w_1)/(\partial\hat{u}_3/\partial w_3), \tag{24}$$

which, when set greater than zero and when amended by the substitution of symmetry relations, yields the necessary and sufficient condition that

$$\frac{\partial\hat{u}_1/\partial w_3}{\partial\hat{u}_3/\partial w_3} > \frac{\partial\hat{u}_{21}/\partial w_3}{\partial\hat{u}_{23}/\partial w_3}, \tag{25}$$

or equivalently, that

$$\frac{\partial\hat{u}_{11}/\partial w_3 + \partial\hat{u}_{21}/\partial w_3}{\partial\hat{u}_{13}/\partial w_3 + \partial\hat{u}_{23}/\partial w_3} > \frac{\partial\hat{u}_{21}/\partial w_3}{\partial\hat{u}_{23}/\partial w_3}. \tag{26}$$

Simple manipulation of the inequality (26)[4] reveals that it holds if and only if

$$\frac{\partial \hat{u}_{11}/\partial w_3}{\partial \hat{u}_{13}/\partial w_3} > \frac{\partial \hat{u}_{21}/\partial w_3}{\partial \hat{u}_{23}/\partial w_3}. \tag{27}$$

Thus, the condition (27) is necessary and sufficient for the transfer of the rationed input from the first to the second line. Since the ratios are negative, the change in the first input relative to the third must be larger in absolute value in the second production line than in the first when the price of the rationed input varies Thus, the transfer takes advantage of the second line's greater ability to economize on the first input.

The final step in the plausibility argument requires the condition for the transfer be linked to the occurrence of a positive substitution effect in the first line. By returning to (21) and solving for the requirement that $\partial u_{11}/\partial w_1 > 0$, the condition

$$\frac{\partial \hat{u}_{11}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1} > \frac{\partial \hat{u}_1/\partial w_3}{\partial \hat{u}_3/\partial w_3}, \tag{28}$$

or equivalently,

$$\frac{\partial \hat{u}_{11}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1} > \frac{\partial \hat{u}_{11}/\partial w_3 + \partial \hat{u}_{21}/\partial w_3}{\partial \hat{u}_{13}/\partial w_3 + \partial \hat{u}_{23}/\partial w_3}, \tag{29}$$

is found to be necessary and sufficient when $\partial \hat{u}_{i3}/\partial w_1 > 0$ for $i = 1, 2$.

The two conditions (27) and (28), one necessary and sufficient for the transfer and the other for the positive substitution effect, can be related by introducing three inequalities:

$$\frac{\partial \hat{u}_{i1}/\partial w_1}{\partial \hat{u}_{i3}/\partial w_1} < \frac{\partial \hat{u}_{i1}/\partial w_3}{\partial \hat{u}_{i3}/\partial w_3}, \quad i = 1, 2 \quad \text{and} \quad \partial \hat{u}_{i3}/\partial w_1 > 0, \tag{30}$$

which results from the concavity of C_i in $\bar{w}$; and

$$\frac{\partial \hat{u}_1/\partial w_1}{\partial \hat{u}_3/\partial w_1} < \frac{\partial \hat{u}_1/\partial w_3}{\partial \hat{u}_3/\partial w_3} \quad \text{for} \quad \partial \hat{u}_3/\partial w_1 > 0, \tag{31}$$

which follows from the concavity of $\sum C_i$ in $\bar{w}$. If it assumed that $\partial u_{11}/\partial w_1 > 0$, then from (30) and (29), it is clear that

$$\frac{\partial \hat{u}_{11}/\partial w_3}{\partial \hat{u}_{13}/\partial w_3} > \frac{\partial \hat{u}_{11}/\partial w_3 + \partial \hat{u}_{21}/\partial w_3}{\partial \hat{u}_{13}/\partial w_3 + \partial \hat{u}_{23}/\partial w_3}, \tag{32}$$

---

[4] Let the elements of the inequality (26) be represented respectively by $(a_1 + b_1)/(a_2 + b_2) > b_1/b_2$, where $a_2, b_2 < 0$ and $a_1, b_1 > 0$. Multiplying both sides by $a_2 + b_2$ and subtracting b_1, $a_1 < a_2 b_1/b_2$, which is equivalent to $a_1/a_2 > b_1/b_2$, the inequality (27). Since (27) and (26) are logically equivalent, the condition (27) is necessary and sufficient for the transfer.

which is true if and only if[5]

$$\frac{\partial \hat{u}_{11}/\partial w_3}{\partial \hat{u}_{13}/\partial w_3} > \frac{\partial \hat{u}_{21}/\partial w_3}{\partial \hat{u}_{23}/\partial w_3}. \tag{33}$$

Therefore, since (33) and (27) are identical, a positive substitution effect in the first line implies a transfer of the rationed input to the second line. Although the transfer is necessary, it is not sufficient to reverse the sign of the substitution effect since (32) does not imply (29). Thus, the final two suppositions of the plausibility argument are logically verified.

For the sake of completeness, it should be noted that (33) is only a part of a more restrictive necessary condition, which can be derived from (29) and the concavity relations. First, from (31) and (29)

$$\frac{\partial \hat{u}_{11}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1} > \frac{\partial \hat{u}_{11}/\partial w_1 + \partial \hat{u}_{21}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1 + \partial \hat{u}_{23}/\partial w_1}, \tag{34}$$

for which it is necessary and sufficient that[6]

$$\frac{\partial \hat{u}_{11}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1} > \frac{\partial \hat{u}_{21}/\partial w_1}{\partial \hat{u}_{23}/\partial w_1}. \tag{35}$$

However, since (34) is not sufficient to imply (29), (35) is no more than a necessary condition for the positive substitution effect. Its interpretation is analogous to that of (27) where, in this case, the ratios of change are compared for a variation in the price of the first input and found to be greater in absolute value in the second line than in the first. Second, from the logical equivalence of (32), (33), and (26), it also follows from (29) that

$$\frac{\partial \hat{u}_{11}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1} > \frac{\partial \hat{u}_{21}/\partial w_3}{\partial \hat{u}_{23}/\partial w_3}. \tag{36}$$

Thus, these three conditions in (36), (35), and (33), actually reduce to one necessary condition that

$$\frac{\partial \hat{u}_{11}/\partial w_3}{\partial \hat{u}_{13}/\partial w_3} > \frac{\partial \hat{u}_{11}/\partial w_1}{\partial \hat{u}_{13}/\partial w_1} > \frac{\partial \hat{u}_{21}/\partial w_3}{\partial \hat{u}_{23}/\partial w_3} > \frac{\partial \hat{u}_{21}/\partial w_1}{\partial \hat{u}_{23}/\partial w_1}, \tag{37}$$

---

[5] Let the elements of (32) be represented respectively by $a_1/a_2 > (a_1 + b_1)/(a_2 + b_2)$, where $a_2, b_2 < 0$ and $a_1, b_1 > 0$. Multiplying by $a_2 + b_2$ and subtracting a_1, $a_1 b_2/a_2 < b_1$ so that $a_1/a_2 > b_1/b_2$. Since the logic can be reversed, the two inequalities are logically equivalent.

[6] Let the elements in (34) be represented respectively by $c_1/c_2 > (c_1 + d_1)/(c_2 + d_2)$, where $c_1, d_1 < 0$ and $c_2, d_2 > 0$. Multiplying by $c_2 + d_2$ and subtracting c_1, $c_1 d_2/c_2 > d_1$, which is equivalent to $c_1/c_2 > d_1/d_2$.

which is not sufficient to imply (28), the positive substitution effect, but does imply that the rationed input must be transferred from the first to the second production line.

V. OUTPUT-SPECIFIC SUBSTITUTION EFFECTS IN THE PRESENCE OF A JOINT INPUT

Turning next to the case of joint or "public" inputs, that is, $x_{13} = x_{23} = x_3$, it will be shown that the perverse substitution effect can occur only when the first and third inputs are complementary in one line and competitive in the other. Minimizing cost subject to the well-behaved production functions $y_i = f_i(x_{i1}, x_{i2}, x_3)$, $i = 1, 2$, results in the constant-output demand functions $\hat{u}_{ij} = \hat{u}_{ij}(\bar{w}, \bar{y})$, $j = 1, 2$, and $\hat{u}_3 = \hat{u}_3(\bar{w}, \bar{y})$ and the twice continuously differentiable, *nonseparable* cost function, $C = C(\bar{w}, \bar{y})$, concave in $\bar{w}$.

Fixing the value of the joint input at its optimum, the conditional demands can be expressed as

$$\hat{u}_{ij} = u_{ij}(w_1, w_2, \bar{y}, \hat{u}_3(\bar{w}, \bar{y})). \tag{38}$$

Letting $i, j = 1$ and differentiating with respect to w_1, the specific substitution effect is found to be

$$\begin{aligned} \partial\hat{u}_{11}/\partial w_1 &= \partial u_{11}/\partial w_1 + (\partial u_{11}/\partial\hat{u}_3)(\partial\hat{u}_3/\partial w_1) \\ &= \partial u_{11}/\partial w_1 + (\partial u_{11}/\partial\hat{u}_3)(\partial\hat{u}_{11}/\partial w_3 + \partial\hat{u}_{21}/\partial w_3), \end{aligned} \tag{39}$$

using the substitution

$$\begin{aligned} \partial\hat{u}_3/\partial w_1 &= \partial^2 C/\partial w_1\,\partial w_3 = \partial^2 C/\partial w_3\,\partial w_1 \\ &= \partial\hat{u}_1/\partial w_3 = \partial\hat{u}_{11}/\partial w_3 + \partial\hat{u}_{21}/\partial w_3. \end{aligned} \tag{40}$$

Noting that $\partial\hat{u}_{11}/\partial w_3 = (\partial u_{11}/\partial\hat{u}_3)(\partial\hat{u}_3/\partial w_3)$, solving for $\partial u_{11}/\partial\hat{u}_3$, and substituting it into $\partial\hat{u}_{11}/\partial w_1$,

$$\begin{aligned} \partial\hat{u}_{11}/\partial w_1 = \partial u_{11}/\partial w_1 &+ (\partial\hat{u}_{11}/\partial w_3)^2/\partial\hat{u}_3/\partial w_3 \\ &+ (\partial\hat{u}_{11}/\partial w_3)(\partial\hat{u}_{21}/\partial w_3)/\partial\hat{u}_3/\partial w_3, \end{aligned} \tag{41}$$

Since for a fixed value of x_3, the variable cost function is separable in $\bar{y}$ and the individual $C_i(\bar{w}, \bar{y}, x_3)$ is concave in $\bar{w}$, $\partial u_{11}/\partial w_1 \leqq 0$. By the concavity of $C(\bar{w}, \bar{y})$ in $\bar{w}$, $\partial\hat{u}_3/\partial w_3 \leq 0$ so the second term of the right member is also negative. However, the third term calls into question the a priori negativity of the specific substitution effect when the first and third inputs are competitive in one line and complementary in the other, the converse of the case of rationing.

Finally, it should be noted that the "public" input may be considered an input employed specifically in one line but which creates "external" effects in the other, as in the case of bees employed in the production of honey, but which pollinate apple blossoms in a neighboring orchard.

VI. A NUMERICAL EXAMPLE

The goal of the argument so far has been to demonstrate the logical possibility of perverse substitution effects. In this section a numerical example of the rationing case will be presented.[7] It will be constructed from the generalized Leontief cost function (GLCF),

$$C_i = y_i \sum_{j=1}^{3} \sum_{k=1}^{3} a^i_{jk} w_j^{1/2} w_k^{1/2}, \qquad a^i_{jk} = a^i_{kj},\ a^i_{jk} \gtrless 0, \tag{42}$$

whose parameters are selected so that, for the given values of factor prices and the rate of output, it is a well-behaved cost function.[8]

To ensure the cost function's monotonicity, the choice of parameters for the given $(\bar{w}, y_i)$ must result in nonnegative values of the factor demands. From Shephard's lemma, the factor demands are

$$\hat{u}_{ij} = \partial C_i/\partial w_j = y_i \sum_{k=1}^{3} a^i_{jk} (w_k/w_j)^{1/2}. \tag{43}$$

Moreover, the GLCF must be concave in $\bar{w}$ at the given values $(\bar{w}, y_i)$, which requires its Hessian matrix to exhibit negative semidefinite form. The elements of the Hessian are

$$\partial \hat{u}_{ij}/\partial w_k = \partial^2 C_i/\partial w_k\, \partial w_j = (1/2) y_i a^i_{jk} w_j^{-1/2} w_k^{-1/2} \qquad \text{for} \quad j \neq k \tag{44}$$

and

$$\partial \hat{u}_{ij}/\partial w_j = \partial^2 C_i/\partial w_j^2 = (-1/2) y_i \sum_{\substack{k=1 \\ k \neq j}}^{3} a^i_{jk} w_j^{-3/2} w_k^{1/2}. \tag{45}$$

Finally, an inspection of the GLCF shows that the requirement of linear homogeneity in $\bar{w}$ is fulfilled by any vector $\bar{w}$.

The GLCF in this example is evaluated at $y_1 = 2$, $y_2 = 2$,

---

[7] The author is grateful to Shaun Bamford, Gary Gigliotti, Barry Kolatch, and Robert Moffitt for their helpful discussion on functional forms for cost.

[8] See Diewert [2] for a discussion of the GLCF and the requirements it must satisfy.

$w_1 = w_2 = w_3 = 1$. For these values, the following coefficients satisfy the requirements on the cost function:

$$
\begin{aligned}
&a^1_{11} = 3, \quad && a^1_{12} = a^1_{21} = -2, \quad && a^1_{13} = a^1_{31} = 5,\\
&a^1_{22} = 1, && a^1_{23} = a^1_{32} = 4,\\
&a^1_{33} = 3,\\
&a^2_{11} = 4, && a^2_{12} = a^2_{21} = 1, && a^2_{13} = a^2_{31} = 9,\\
&a^2_{22} = 6, && a^2_{23} = a^2_{32} = 1,\\
&a^2_{33} = 5.
\end{aligned}
$$

The values of the demands are $\hat{u}_1 = 12$, $\hat{u}_{12} = 6$, $\hat{u}_{13} = 24$, $\hat{u}_{21} = 28$, $\hat{u}_{22} = 16$, $\hat{u}_{23} = 30$, and the Hessian matrices take the form

$$
\begin{bmatrix}
\partial\hat{u}_{i1}/\partial w_1 & \partial\hat{u}_{i1}/\partial w_2 & \partial\hat{u}_{i1}/\partial w_3\\
\partial\hat{u}_{i2}/\partial w_1 & \partial\hat{u}_{i2}/\partial w_2 & \partial\hat{u}_{i2}/\partial w_3\\
\partial\hat{u}_{i3}/\partial w_1 & \partial\hat{u}_{i3}/\partial w_2 & \partial\hat{u}_{i3}/\partial w_3
\end{bmatrix}
=
\begin{bmatrix}
-(a^i_{12} + a^i_{13}) & a^i_{12} & a^i_{13}\\
a^i_{21} & -(a^i_{21} + a^i_{23}) & a^i_{23}\\
a^i_{31} & a^i_{32} & -(a^i_{32} + a^i_{31})
\end{bmatrix},
\tag{46}
$$

whose values are

$$
\begin{bmatrix}
-3 & -2 & 5\\
-2 & -2 & 4\\
5 & 4 & -9
\end{bmatrix}
\quad \text{for} \quad i = 1
\tag{47}
$$

and

$$
\begin{bmatrix}
-10 & 1 & 9\\
1 & -2 & 1\\
9 & 1 & -10
\end{bmatrix}
\quad \text{for} \quad i = 2.
\tag{48}
$$

Both matrices exhibit negative semidefinite form. (The singularity of the third-order principal minor is due to the linear homogeneity of the cost function in $\bar{w}$.)

The values of the output-specific substitution effects under rationing are determined by substituting these values in (47) and (48) into (21):

$$
\begin{aligned}
\partial u_{21}/\partial w_1 &= -10 + (126/19)\\
&= -64/19
\end{aligned}
\tag{49}
$$

and

$$
\begin{aligned}
\partial u_{11}/\partial w_1 &= -3 + (70/19)\\
&= 13/19
\end{aligned}
\tag{50}
$$

Although the substitution effect is positive in the first line, the aggregate effect is negative: $\partial(u_{11} + u_{21})/\partial w_1 = -51/19$.

Generalizing the expression in (24),

$$\partial u_{i3}/\partial w_1 = \partial \hat{u}_{i3}/\partial w_1 - (\partial \hat{u}_{i3} \partial w_3)(\partial \hat{u}_3/\partial w_1)/(\partial \hat{u}_3/\partial w_3). \tag{51}$$

Substituting its numerical values, $\partial u_{13}/\partial w_1 = -31/19$ and $\partial u_{23}/\partial w_1 = 31/19$, which supports the assertion in Section IV that the reduction in u_{21} requires an increase in the competitive u_{23} and forces a reduction of u_{13} and an increase of u_{11}. A glance at the appropriate elements, 5 and 9, of the Hessian matrices reveals that the first and third inputs are competitive in both lines.

VII. IMPLICATIONS FOR PUBLIC POLICY

The use of taxes and subsidies is often suggested to encourage or discourage the designs of economic agents. One critical assumption underlying these schemes is that demand is inversely related to price, but the foregoing examination of the theory of multiple production has questioned its validity in the case of an unconstrained, profit-maximizing firm's demand for an input in a particular production process.

For example, to encourage labor's employment, the government might institute a wage subsidy for unskilled labor. The theory predicts that multiproduct firms would increase their overall employment of labor, but employment in the production of certain outputs might actually decrease. Thus, where labor is not perfectly mobile, pockets of unemployment could develop in those industries.

To discourage the use of an environmentally harmful pesticide in farming, the government might impose a tax on it. The theory predicts that farms overall would reduce their application of the chemical; however, they might increase its use on certain crops. In cases where the harm from the limited increase in use outweighs the benefit from the overall reduction, the tax policy is clearly undesirable.[9] Thus, care must be exercised when the distribution of the total effect is not a matter of indifference.

ACKNOWLEDGMENT

The author is grateful to an anonymous Associate Editor for invaluable criticism on earlier drafts.

---

[9] This illustration was suggested by M. B. Stewart.

REFERENCES

1. D. V. T. BEAR, Inferior inputs and the theory of the firm, *J. Pol. Econ.* **73** (1965), 287–289.
2. W. E. DIEWERT, An applicaton of the Shephard duality theorem: A generalized Leontief production function, *J. Pol. Econ.* **79** (1971), 481–507.
3. J. P. HUGHES, Factor demand in the multi-product firm, *Southern Econ. J.* **45** (1978), 494–501.
4. A. MINGUET, Substitution and output effects in the neoclassical theory of the firm under imperfect competition, *European Econ. Rev.* **6** (1975), 39–57.
5. R. W. PFOUTS, The theory of cost and production in the multi-product firm, *Econometrica* **29** (1961), 650–658.
6. R. A. POLLAK, Conditional demand functions and consumption theory, *Quart. J. Econ.* **83** (1969), 60–78.
7. Y. SAKAI, Substitution and expansion effects in production theory: The case of joint production, *J. Econ. Theory* **9** (1974), 255–274.
8. H. R. VARIAN, "Microeconomic Analysis," Norton, New York, 1978.

JOURNAL OF ECONOMIC THEORY 25, 302–313 (1981)

Stable Quantities in Fixed Price Disequilibrium

JOHN C. ECKALBAR

Department of Economics, California State University, Chico, California 95929

Received August 11, 1979; revised April 25, 1980

I

Recent work in the area of non-Walrasian process analysis has focused on three primary deficiencies in the Walrasian tatonnement assumptions. These are the Walrasian neglect of endowment and spillover effects together with the failure to assign a special role for money. In detail:

First, it has long been known that if trade takes place during the process of price adjustment, the resulting capital gains and losses coupled with the change in the distribution of commodities will tend to shift the position of the equilibrium. (For an early source, see [8, Book V, Chap. II].) This leaves the system with the problem of trying, so to speak, to close in on a moving target. This is the endowment effect. Surprisingly, Hahn, Negishi, and others (see [7; 1, Chap. 13]) have found that stability follows in this class of models more readily than in Walrasian models.

Second, Clower [3] and Patikin [9, Chap. XIII] have argued that if trade takes place away from equilibrium, we should expect that non-zero excess demands in some markets will influence trading plans in other markets. This is the spillover effect. The classic case is that if a worker on the long side of the labor market cutting his planned consumption below its Walrasian level. Because of spillovers, the demands which actually get expressed in the marketplace can differ significantly, even in sign, from the demands which the Walrasian auctioneer responds to. Initially there was speculation that spillovers would generate instability and/or unemployment equilibria; but thus far the theoretical results are mixed (compare, for example, [11, 10]).

Finally, Clower [4] has argued that the Patikinesque macro-money–general equilibrium models have not succeeded in formally distinguishing between money and other durables as a means of payment. Noting that the Walrasian budget constraint implicitly allows trading in any pair of commodities, he urges that a tighter restriction be introduced in the interest of enforcing the rule—"Money buys goods, and goods buy money;

0022-0531/81/050302-12$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

but goods do not buy goods" [4, pp. 207–208]. Work on the stability of systems having a well defined medium of exchange is proceeding slowly, but some results are available from Arrow and Hahn, Howitt, and Eckalbar [1, Chap. 13; 5].

Models which take cognizance of all three of these problems are rare to non-existent. In the present paper we investigate the stability properties of a model with endowment effects, spillovers, and medium of exchange constraints.

II

We imagine a system with two traders, or suitably constructed aggregates—a household and a firm. There are three goods—labor, commodities, and money. Money wages (w) and prices (p) are assumed to be fixed at $\bar{w}$ and $\bar{p}$. A fixed price system is interesting for several reasons: First, since governments sometimes impose wage and price controls, there is practical interest in the question of the stability of non-price adjustments under such a regime. Second, if we believe that prices adjust more slowly than expectations and money distributions, this model may be interpreted as a short run tool. Finally, the study illustrates the use of some efficient methods for the exploration of non-Walrasian systems and is for that reason of value as a purely theoretical exercise.

On Monday morning the firm and the household meet in the labor market. There, the firm's actual demand for labor, L^d, is found by taking the minimum of three quantities: $\tilde{L}^d(\bar{w}/\bar{p})$, $F^{-1}(e)$, and $\bar{m}^f/\bar{w}$. $\tilde{L}^d(\bar{w}/\bar{p})$ is the unconstrained profit maximizing labor demand. It is the firm's aswer to the question, How much labor would you hire if you believed that you would have no trouble selling output and if you were not constrained to pay cash for labor? $F^{-1}(e)$ is the answer to the question, How much labor is required to produce e, the expected commodity demand? ($F(\ )$ is the production function.) And $\bar{m}^f/\bar{w}$ is the answer to question, What is the maximum amount of labor you can hire, given the cash on demand constraint? Where $\bar{m}^f$ is the firm's beginning period money holding. Then the firm's actual demand for labor is $L^d = \min(\tilde{L}^d(\bar{w}/\bar{p}), F^{-1}(e), \bar{m}^f/\bar{w})$.

We assume that the household supplies $\bar{L}^s$, a constant; and in keeping with the Keynesian flavor of this class of models, we further suppose that $\bar{L}^s > L^d$. The actual quantity traded in the labor market is then $L = \min(L^d, \bar{L}^s) = L^d$, and the firm is assumed to pay the household $w \cdot L^d$ units of money at the close of the labor market.

Traders then march over the commodity market, where the firm offers to sell $F(L)$, the whole of its current production. Households attempt to by

$c^d(\ )$, which we suppose takes the specific form $c^d(\bar{m} - \bar{m}^f + \bar{w} \cdot L)$, where $\bar{m}$ is the total supply of money for the system. This seems quite reasonable under the circumstances. $\bar{m} - \bar{m}^f + \bar{w} \cdot L$ is the amount of money the household enters the commodity market with. $c^d(\ )$ then simply determines how much of that money gets spent and how much demanded as ending money holdings. Note that a low demand for labor spills over into commodity demand via the argument L in $c^d(\ )$.

Under the usual assumptions on utility functions, the household will want to demand a positive level of ending money balances. This will ensure that $c^d(\ ) < (\bar{m} - \bar{m}^f + \bar{w} \cdot L)/\bar{p}$, so we need not explicitly impose a cash constraint on the household, though we should understand that the latter is implied.

The quantity traded in the commodity market is $c = \min(F(L), c^d(\ ))$; and as the commodity market closes, the household pays $\bar{p} \cdot c$ units of money to the firm.

Then at the close of the market, dividends are paid and expectations adjusted in light of the actual demand for commodities.

We will suppose that the firm feels two distinct sorts of pressures from the households. First, it feels it must try to maximize expected profits. And second, it should pay out all "surplus money balances" to the household. That is, the household, as owner of the firm, wants the firm to hold working balances only, idle balances are to be transferred to the household. Thus we give the dividend function, $D(\ )$, the following form:

$$\begin{aligned} D(t) &= \pi(t) - |\bar{w} \cdot \min(F^{-1}(e(t)), \tilde{L}^d) - \bar{m}^f(t)| \\ &\qquad \text{if } \bar{m}^f(t) < \bar{w} \cdot \min(F^{-1}(e(t)), \tilde{L}^d), \\ &= \pi(t) \quad \text{if } \bar{m}^f(t) = \bar{w} \cdot \min(F^{-1}(e(t)), \tilde{L}^d), \\ &= \pi(t) + |\bar{m}^f(t) - \bar{w} \cdot \min(F^{-1}(e(t)), \tilde{L}^d)| \\ &\qquad \text{if } \bar{m}^f(t) > \bar{w} \cdot \min(F^{-1}(e(t)), \tilde{L}^d). \end{aligned}$$

In plain language, the first line says that if the firm does not have enough money to hire the expected profit maximizing quantity of labor, $\min(F^{-1}(e), \tilde{L}^d)$, the firm will hold back on its dividend payments, paying only its profits *minus* its monetary deficiency, $\bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d) - \bar{m}^f$. This is *the result of the firm's effort to maximize profits in a world where all* purchases must be for cash. The third line in effect has the firm declaring a special dividend (i.e., in excess of profits) whenever it finds that it is holding more money than necessary for its labor purchases. Finally, $D(t) = \pi(t)$ whenever the firm finds that it has just enough money for its labor purchases.

It is easy to see that $D(\ )$ reduces to

$$D(\ ) = \pi + \bar{m}^f - \bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d).$$

The change in the firm's initial money holdings over time, $\dot{\bar{m}}^f$, will be $\pi - D$, so $\dot{\bar{m}}^f = \bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d(\bar{w}/\bar{p})) - \bar{m}^f$. The other dynamic variable, $\dot{e}$, is assumed to be determined by $c^d(\ ) - e$. Hence, we have the following dynamic system:

$$\begin{aligned} \dot{\bar{m}}^f &= \bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d) - \bar{m}^f, \\ \dot{e} &= c^d(\bar{m} - \bar{m}^f + \bar{w} \cdot \min(\tilde{L}^d, F^{-1}(e), \bar{m}^f/\bar{w})) - e. \end{aligned} \tag{A}$$

We add the following assumptions:

Assumption F. *The function $F(\ )$ is twice differentiable, with $F'(\ ) > 0$, $F''(\ ) < 0$, and $F(0) = 0$.*

Assumption M. $\bar{m} \geqslant \bar{w}\tilde{L}^d(\bar{w}/\bar{p})$.

Assumption M rules out conditions of severe money shortage, since in effect it says that there is always enough money in the economy that if the firm had it *all*, it could afford its national labor demand. The assumption implies that whenever $\bar{m}^f = \bar{m}$, $\dot{\bar{m}}^f \leqslant 0$—so $\bar{m}^f$ is, as it should be, bounded by $\bar{m}$.[1]

Assumption C. *The function c^d is differentiable, with $c^{d\prime}(\ ) > 0$. And further, we assume $\bar{p} \cdot c^{d\prime}(\ ) < 1$.*

The last line in Assumption C is equivalent to Keynes MPC < 1.

System (A) incorporates many of the more desirable features of this class of models:

(i) We have spillovers working on virtually all demands and supplies, e.g., consumption demand depends upon actual labor sales, labor demand depend upon expected commodity demand, etc.

(ii) Money enters as a well-defined medium of exchange.

(iii) Endowment effects follow from the redistribution of money holdings during trading.

Equilibria

We should be careful to note that an equilibrium to (A) has $\dot{\bar{m}}^f = \dot{e} = 0$, but since w and p are fixed at non-Walrasian values, we will not in general expect market clearing at the equilibrium. An equilibrium to (A) is simply a level of expectations and a distribution of money holdings which is dynamically stationary.

---

[1] Thanks to an anonymous referee for raising this issue. In Footnote 3 I briefly discuss the implications of relaxing Assumption M.

Two sorts of equilibria and possible for (A). One has

$$\begin{gathered}\min(F^{-1}(e), \tilde{L}^d) = \tilde{L}^d = \bar{m}^f/\bar{w},\\ c^d(\bar{m}) = e,\end{gathered} \tag{a}$$

and the other has

$$\begin{gathered}\min(F^{-1}(e), \tilde{L}^d) = F^{-1}(e) = \bar{m}^f/\bar{w},\\ c^d(\bar{m}) = e.\end{gathered} \tag{b}$$

Whether (A) has an equilibrium of type (a) or (b) depends upon the parameters of the model. We will have more to say about this shortly.

Stability

We can begin by organizing the system graphically around its equilibrium loci curves. The set giving $\dot{\bar{m}}^f = 0$ is the union of the following sets

$$\begin{gathered}\{(\bar{m}^f, e) \mid \tilde{L}^d = \bar{m}^f/\bar{w} \text{ and } \tilde{L}^d \leqslant F^{-1}(e)\},\\ \{(\bar{m}^f, e) \mid F^{-1}(e) = \bar{m}^f/\bar{w} \text{ and } \tilde{L}^d > F^{-1}(e)\}.\end{gathered}$$

As a preliminary, we draw the lines $\tilde{L}^d = \bar{m}^f/\bar{w}$ and $F^{-1}(e) = \bar{m}^f/\bar{w}$ in Fig. 1. Since $\tilde{L}^d$ is fixed by virtue of our holding w and p constant, the locus $\tilde{L}^d = \bar{m}^f/\bar{w}$ is simply a vertical line. The line $F^{-1}(e) = \bar{m}^f/\bar{w}$ runs out of the origin and has a slope given by $F'(L)/\bar{w}$, which is positive and decreasing.

To get the locus $\dot{\bar{m}}^f = 0$, we need only partition the phase space into the regions where $F^{-1}(e) \gtreqless \tilde{L}^d$. This is easily done by drawing the line $F^{-1}(e) = \tilde{L}^d$ and noting that above this line we have $F^{-1}(e) > \tilde{L}^d$ and below it we have $F^{-1}(e) < \tilde{L}^d$. The locus $\dot{\bar{m}}^f = 0$ is now seen to be the solid kinked curve in Fig. 2.

The curve $\dot{\bar{m}}^f = 0$ together with the dashed line $\tilde{L}^d = F^{-1}(e)$ partition the phase space into the four regions labeled I, II, III, and IV. In region I, $L^d = \min(\bar{m}^f/\bar{w}, F^{-1}(e), \tilde{L}^d) = \bar{m}^f/\bar{w}$ and $\tilde{L}^d = \min(\tilde{L}^d, F^{-1}(e))$. This is checked by noting that left of the line $\tilde{L}^d = \bar{m}^f/\bar{w}$ we have $\tilde{L}^d > \bar{m}^f/\bar{w}$, and above the line $F^{-1}(e) = \bar{m}^f/\bar{w}$ we have $F^{-1}(e) > \bar{m}^f/\bar{w}$, while above $\tilde{L}^d = F^{-1}(e)$, $\tilde{L}^d < F^{-1}(e)$. In region II, $L^d = \tilde{L}^d$ and $\tilde{L}^d = \min(\tilde{L}^d, F^{-1}(e))$; in III, $L^d = F^{-1}(e)$ and $F^{-1}(e) = \min(\tilde{L}^d, F^{-1}(e))$; and in IV, $L^d = \bar{m}^f/\bar{w}$ and $F^{-1}(e) = \min(\tilde{L}^d, F^{-1}(e))$.

Looking back to the equation for $\dot{\bar{m}}^f$ in (A), it is now easy to verify that northwest of the locus $\dot{\bar{m}}^f = 0$ we have $\dot{\bar{m}}^f > 0$, while $\dot{\bar{m}} < 0$ in I and II.

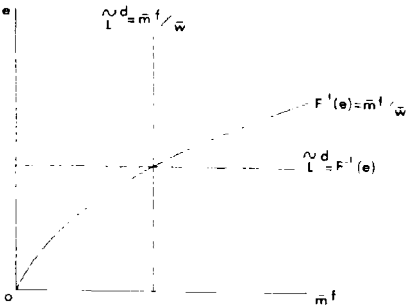

e
$\tilde{L}^d = \bar{m}^f/\bar{w}$
$F^{-1}(e) = \bar{m}^f/\bar{w}$
$\tilde{L}^d = F^{-1}(e)$
o
$\bar{m}^f$

FIGURE 1

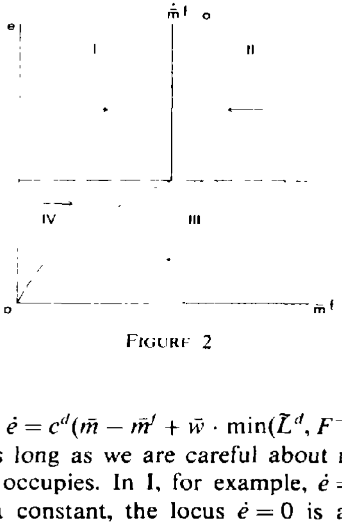


FIGURE 2

Drawing the locus $\dot{e} = c^d(\bar{m} - \bar{m}^f + \bar{w} \cdot \min(\tilde{L}^d, F^{-1}(e), \bar{m}^f/\bar{w}) - e = 0$ is now straightforward as long as we are careful about noting which sector of the phase space $\dot{e} = 0$ occupies. In I, for example, $\dot{e} = 0$ implies $c^d(\bar{m}) = e$. Then since $c^d(\bar{m})$ is a constant, the locus $\dot{e} = 0$ is a horizontal line in I. Intuitively, the firm is cash constrained and is spending the whole of its money balances in I, so the household enters the commodity market holding all of the system's money supply. Then everywhere in I the household's commodity demand is the same.

Above the locus $\dot{e} = 0$ in I, $c^d(\bar{m}) < e$ and therefore $\dot{e} < 0$, and below $\dot{e} = 0$, $\dot{e} > 0$.[2]

If $\dot{e} = 0$ runs through sector II, it will be located by the equation $c^d(\bar{m} - \bar{m}^f + \bar{w} \cdot \tilde{L}^d) = e$. The slope of this line is simply $-c^{d\prime}(\ ) < 0$. Again, above the locus we have $c^d(\ ) < e$ and $\dot{e} < 0$, and vice versa below the locus.

---

[2] The properties of $\dot{e}$ are the same if $\dot{e} = 0$ runs through sector IV rather than sector I.

In sector III, $\dot{e} = 0$ is given by $c^d(\bar{m} - \bar{m}^f + \bar{w} \cdot F^{-1}(e)) = e$. The slope of this line is

$$\frac{de}{d\bar{m}^f} = \frac{c^{d\prime}(\ )}{\dfrac{c^{d\prime}(\ )\,\bar{w}}{F'(L)} - 1}.$$

Recall that in III, $\min(\tilde{L}^d, \bar{m}^f/\bar{w}, F^{-1}(e)) = F^{-1}(e)$. Since the firm's labor demand is expected-sales constrained, elementary price theory implies $F'(L) > \bar{w}/\bar{p}$. Then $\bar{p} \cdot c^{d\prime}(\ ) < 1$ from Assumption C implies $c^{d\prime}(\ ) \cdot \bar{w} < F'(L)$, and the slope $de/d\bar{m}^f < 0$ in III. Note again that above this locus we have $c^d(\ ) < e$ and $\dot{e} < 0$, and vice versa.

Since (A) can have two sorts of equilibria, we must consider two stability problems. Figure 3 shows (A) configured such that there is a "notional labor demand constrained" equilibrium at point a with

$$L^d = \min(\tilde{L}^d, F^{-1}(e), \bar{m}^f/\bar{w}) = \tilde{L}^d = \bar{m}^f/\bar{w},$$
$$c^d(\bar{m}) = e.$$

It should be clear from the picture that point a is stable. Notice also that there are three "sliding lines" (see [2, p. 122]) leading to a, i.e., once a trajectory is carried onto, say $\dot{\bar{m}}^f = 0$, it then slides along that line to a with no further motion in the dimension m.

Figure 4 is drawn with the second sort of equilibrium. At b we have

$$L^d = \min(\tilde{L}^d, F^{-1}(e), \bar{m}^f/\bar{w}) = F^{-1}(e) = \bar{m}^f/\bar{w},$$
$$c^d(\bar{m}) = e.$$

Again, we see from the picture that b is a stable equilibrium with a sliding from the left.[3]

Though the stability of points a or b is rather obvious from the pictures, it is worth our while to establish it formally, since the effort will reveal much

---

[3] If we give up Assumption M, then to ensure that $\bar{m}^f \not> \bar{m}$, the first line in (A) will have to be replaced by

$$\begin{aligned} \dot{\bar{m}}^f &= \bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d) - \bar{m}^f, && \text{unless } \bar{m}^f = \bar{m} \text{ and } \bar{m} < \bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d) \\ &= 0, && \text{if } \bar{m}^f = \bar{m} \text{ and } \bar{m} < \bar{w} \cdot \min(F^{-1}(e), \tilde{L}^d). \end{aligned}$$

We cannot explore this situation in detail, but a quick look at the graphs should establish that this creates no special problem. If $\bar{m} < \bar{w}\tilde{L}^d$, the locus $\{(e, \bar{m}^f) \mid \dot{\bar{m}}^f = 0\}$ will have its vertical section above the point where $\bar{m}^f = \bar{m}$ and to the left of the line given by $\bar{w}\tilde{L}^d = \bar{m}^f$. The rest of the analysis follows as before, except that the system cannot get to the right of the line $\bar{m}^f = \bar{m}$.

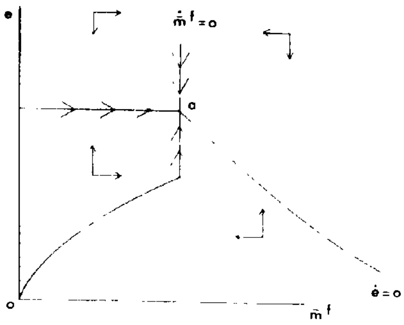


FIGURE 3

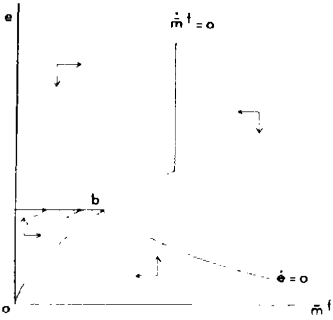


FIGURE 4

about the problems of working with non-Walrasian process. It is sufficient for our purpose to check the stability of point *b* only.

Begin by noting that in the neighborhood of point *b* the dynamic variables e and $\bar{m}^f$ can be driven by two distinct dynamic systems. Below and to the right of *b*, e and $\bar{m}^f$ are driven by

$$\begin{aligned} \dot{\bar{m}}^f &= \bar{w} \cdot F^{-1}(e) - \bar{m}^f, \\ \dot{e} &= c^d(\bar{m} - \bar{m}^f + \bar{w} \cdot F^{-1}(e)) - e, \end{aligned} \qquad (\mathrm{A^{III}})$$

while above and to the left of point *b*, e and $\bar{m}^f$ are moved by

$$\begin{aligned} \dot{\bar{m}}^f &= \bar{w} \cdot F^{-1}(e) - \bar{m}^f, \\ \dot{e} &= c^d(\bar{m}) - e. \end{aligned} \qquad (\mathrm{A^{IV}})$$

The line $\dot{\bar{m}}^f = 0$ is a "switching line," i.e., it is the boundary between the region where e and $\bar{m}^f$ are governed by ($\mathrm{A^{III}}$) and the region where e and $\bar{m}^f$

driven by $(\mathrm{A}^{\mathrm{IV}})$. As a trajectory is pushed across this switching line, it comes under the influence of a different dynamical system. The presence of such switching lines is characteristic of non-Walrasian proceses (see [10, 6]).

Speaking generally, the formal stability problem is this: Two distinct dynamic systems, $(\mathrm{A}^{\mathrm{III}})$ and $(\mathrm{A}^{\mathrm{IV}})$, are defined on the same variables. At any point in the neighborhood of b one of the systems is driving the dynamic variables and the other system is dormant, though the motion of the system may carry the variables to a point where the active and dormant systems reverse roles. In fact, this may happen numerous times as the variables move through the phase space. The question is, under what conditions can such a Jekyll and Hyde system be shown to be stable?

I have dealt with this question at length elsewhere, (see [6]) so in the present paper I will simply take those results as given and illustrate their usefulness on system (A).

The following theorems, drawn from [6], are used in the proof. The theorems are adapted for the present notation, where $(\hat{\bar{m}}^f, \hat{e})$ is the equilibrium.

THEOREM 1. *If* (i) *there exists a continuous positive definite function* $V(\bar{m}^f - \hat{\bar{m}}^f, e - \hat{e})$, *with* (ii) $\dot{V}(\bar{m}^f - \hat{\bar{m}}^f, e - \hat{e}) < 0$ *outside* M, *and* (ii) $\dot{V}(\bar{m}^f - \hat{\bar{m}}^f, e - \hat{e}) \leq 0$ *on* M, *where* (iv) M *is a set containing no entire trajectories other than the null trajectory, then the origin is asymptotically stable.*

Essentially, this is Lyapunov's familiar stability theorem with the addendum that V can be momentarily constant as the trajectory passes through a set M.

Considering the two dimensional linearized system in the variables $\bar{m}^f$ and e,

$$\begin{aligned} \dot{\bar{m}}^f &= a(\bar{m}^f - \hat{\bar{m}}^f) + b(e - \hat{e}), \\ \dot{e} &= c(\bar{m}^f - \hat{\bar{m}}^f) + d(e - \hat{e}), \end{aligned} \tag{L}$$

where a, b, c, and d are constants, we have the following theorem.

THEOREM 2. *If the Routh–Hurwitz conditions* $(a + d < 0$ *and* $ad - bc > 0)$ *hold for* (L), *i.e., if* (L) *is asymptotically stable, then there is a function* $V(\bar{m}^f - \hat{\bar{m}}^f, e - \hat{e})$ *satisfying the conditions of Theorem* 1. *And with* $b \neq 0$, $V(\ ) = [a(\bar{m}^f - \hat{\bar{m}}^f) + b(e - \hat{e})]^2 + (ad - bc)(\bar{m}^f - \hat{\bar{m}}^f)^2$.

Finally, defining $x = (\bar{m}^f - \hat{\bar{m}}^f, e - \hat{e})$, we state the main theorem.

THEOREM 3. *If two dynamic systems,* $f(x)$ *and* $g(x)$, *are defined on the same variables, if these two systems have the same isolated equilibrium, and* *if there is a* single *function $V(\ )$ which is consistent with the assumptions of Theorem* 1 *for both $f(x)$ and $g(x)$ separately, then by hybrid system*

$$\begin{aligned} \dot{x} &= f(x), && x \in S^a, \\ &= g(x), && x \in S^b, \\ &= f(x) = g(x), && x \in S^a \cap S^b, \end{aligned}$$

where S^a and S^b are subsets of the phase space, is asymptotically stable.

Using these theorems, we now prove the following theorem.

THEOREM 4. *The point b is locally asymptotically stable.*

Proof. We first have to show that (A^{III}) and (A^{IV}) have the same equilibrium, $(\hat{\bar{m}}^f, \hat{e})$. Note that any point $(\bar{m}^{f*}, e^*)$ which has $\dot{\bar{m}}^f = 0$ in (A^{III}) also has $\dot{\bar{m}}^f = 0$ in (A^{IV}); and further, if $\dot{\bar{m}}^f = 0$, $F^{-1}(e^*) = \bar{w}/\bar{m}^{f*}$. But then $c^d(\bar{m} - \bar{m}^{f*} + \bar{w}/F^{-1}(e^*)) = c^d(\bar{m})$, so $\dot{e} = 0$ in (A^{III}) iff $\dot{e} = 0$ in (A^{IV}). Hence (A^{III}) and (A^{IV}) share the equilibrium $(\hat{\bar{m}}^f, \hat{e})$.

We now show that there is a function $V(\ )$ which is continuous, positive definite, zero at the equilibrium, and decreasing (or only momentarily constant) along trajectories on *both* sides of the switching line in a neighborhood of the equilibrium.

A linear approximation of (A^{III}) in a neighborhood of the equilibrium is

$$\begin{aligned} \dot{\bar{m}}^f &= -(\bar{m}^f - \hat{\bar{m}}^f) + |\bar{w}/F'(L)| \cdot (e - \hat{e}), \\ \dot{e} &= -c^{d'}(\ ) \cdot (\bar{m} - \hat{\bar{m}}^f) + |(c^{d'}(\ )\bar{w}/F'(L)) - 1| \cdot (c - \hat{e}). \end{aligned} \qquad (A_L^{III})$$

The Jacobian to (A_l^{III}) is

$$J_l^{III} = \begin{pmatrix} -1 & \bar{w}/F'(L) \\ -c^{d'}(\ ) & \dfrac{c^{d'}(\ )\,\bar{w}}{F'(L)} - 1 \end{pmatrix}.$$

Clearly, $\operatorname{tr}(J_L^{III}) < 0$ and $|J_L^{III}| = 1 > 0$. Using Theorem 2, this establishes (see [6]) that the function

$$V^{III}(\ ) = \left[-(\bar{m}^f - \hat{\bar{m}}^f) + \frac{\bar{w}}{F'(L)}(e - \hat{e}) \right]^2 + |J_l^{III}| \cdot (\bar{m}^f - \hat{\bar{m}}^f)^2$$

is decreasing (or only momentarily constant)[4] along all trajectories and satisfies the assumptions of Theorem 1.

---

[4] $\dot{V} = 2(\operatorname{tr}(J_l^{III}))(\dot{\bar{m}}^f)^2$, so $\dot{V} < 0$ except along the locus $\dot{\bar{m}}^f = 0$, where $\dot{V} = 0$. But the locus $\dot{\bar{m}}^f = 0$ contains no entire trajectories other than the null trajectory. This can be checked by looking at Fig. 4 and noting that half trajectories emanating from points on $\dot{\bar{m}}^f$ (other than b) point either straight up or straight down, while the locus $\dot{\bar{m}} = 0$ is positively sloped.

Now we show that the same function is decreasing along trajectories on the other side of the switching line. Linearizing (A^{IV}) we have

$$\begin{aligned} \dot{\bar{m}}^f &= -(\bar{m}^f - \hat{\bar{m}}^f) + [\bar{w}/F'(L)] \cdot (e - \hat{e}), \\ \dot{e} &= -(e - \hat{e}). \end{aligned} \tag{A_L^{IV}}$$

The Jacobian is

$$J_L^{IV} = \begin{pmatrix} -1 & \bar{w}/F'(L) \\ 0 & -1 \end{pmatrix}.$$

And again $\operatorname{tr}(J_L^{IV}) < 0$ and $|J_L^{III}| = 1 > 0$. This establishes that the function $V^{IV}(\ ) = [-(\bar{m}^f - \hat{\bar{m}}^f) + (\bar{w}|F'(L))(e - \hat{e})]^2 + |J_L^{IV}| \cdot (\bar{m}^f - \hat{\bar{m}}^f)^2$ has the desired properties. But $|J_L^{IV}| = |J_L^{III}| = 1$, so the functions $V^{III}(\ )$ and $V^{IV}(\ )$ are identical. Hence, there is a *single* Lyapunov function which is decreasing along trajectories on *both* sides of the switching line, and local asymptotic stability is established for the hybrid system by Theorem 3.

III

We have put together a non-Walrasian model that has many of the features called for in the past few years: trade takes place away from equilibrium with non-zero excess demands spilling over into other markets, *m* is well defined as a medium of exchange, and endowments change. Using tools recently developed explicitly for the study of non-Walrasian processes, the system is found to be stable under some familiar, and rather weak, Keynesian assumptions.

Much remains to be done in the study of non-Walrasian processes. More complex systems should be investigated. Of special interest are questions of existence and stability of increasingly realistic systems with a medium of exchange and other financial assets. Work on stronger stability theorems should also be given a high priority.

Ultimately, this line of inquiry seems to promise movement to a whole new level of understanding of the dynamics of economic systems, and with this a second generation of macro policy discussion.

ACKNOWLEDGMENTS

The author is pleased to thank Neil Raymon and an anonymous referee for their helpful comments.

REFERENCES

1. K. J. ARROW AND F. H. HAHN, "General Competitive Analysis," Holden–Day, San Francisco, 1971.
2. E. A. BARBASHIN, "Introduction to the Theory of Stability," Wolters–Noordhoff, Groningen, 1970.
3. R. W. CLOWER, The Keynesian counter-revolution: A theoretical appraisal, *in* "The Theory of Interest Rates" (Hahn and Brechling, Eds.), pp. 103–125, Macmillan, London, 1965.
4. R. W. CLOWER, A reconsideration of the microfoundations of monetary theory, *Western Econ. J.* **6** (1967), 1–9.
5. J. C. ECKALBAR, Stability with medium of exchange constraints and spillovers, *Oxford Econ. Papers* **31** (1979), 387–402.
6. J. C. ECKALBAR, The stability of non-Walrasian processes, *Econometrica* **48** (1980), 371–386.
7. F. H. HAHN AND T. NEGISHI, A theorem on non-Tatonnement stability, *Econometrica* **30** (1962), 463–469.
8. P. W. HOWITT, Stability and the quantity theory, *J. Pol. Econ.* **82** (1974), 133–151.
9. A. MARSHALL, "Principles of Economics," 9th (variorum) ed., Macmillan, London, 1961.
10. D. PATINKIN, "Money, Interest, and Prices," Harper & Row, New York, 1965.
11. H. VARIAN, Non-Walrasian equilibria, *Econometrica* **45** (1977), 573–590.
12. E. C. H. VEENDORP, Stable spillovers among substitutes, *Rev. Econ. Studies* **131** (1975), 445–456.

JOURNAL OF ECONOMIC THEORY 25, 314–317 (1981)

Approximation of Points of the Convex Hull of a Sum of Sets by Points of the Sum: An Elementary Approach*

ROSS M. STARR

Department of Economics, University of California, San Diego, La Jolla, California 92093

Received November 5, 1979; revised April 13, 1980

The Shapley–Folkman theorem and its corollaries [1, 2, 3, 4, 5, 6, 8] provide strong bounds on the distance between the sum of a family of nonconvex sets and the convex hull of the sum. Proofs of the theorem are nonconstructive, and require moderately advanced analysis. The proof developed below is based on elementary considerations. It provides an approximation sequentially with the successive addition of sets to the sum. The approximation is not so close as that provided by the Shapley–Folkman theorem, but for any given point of the convex hull we will find a specific point in the sum within a previously determined bound on the distance between the two.

Particular virtues of the bounds associated with the Shapley–Folkman theorem are the relatively tight approximation developed and its behavior as the number of summands becomes large. The bounded distance between the sum and its convex hull depends not on the number of sets summed (denoted m) as this number becomes large, but rather on the dimensionality of the space (denoted N). Thus as the number of summands becomes large, the average discrepancy between the sum and its convex hull converges to 0 as $1/m$. The bounds developed below, on the contrary, vary as $m^{1/2}$ so that the average discrepancy converges to 0 as $1/m^{1/2}$. The virtue of the results here is their comparative ease of proof and the sequential construction making it relatively easy to find a point of the sum nearby to any chosen point of the convex hull.

For $S \subset R^N$, S compact, we define several measures of size and nonconvexity. con S denotes the convex hull of S.

---

* A sketch of the analysis of this paper was developed in 1965 in an early draft of [8]. It was not published then, since it seemed to be superseded by the Shapley–Folkman theorem. Subsequent discussions have convinced me of its independent interest, particularly in relation to theorem 1 of [4] and the theorem of [7].

0022-0531/81/050314-04$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

The *radius* of S is

$$\operatorname{rad}(S) \equiv \inf_{x \in R^N} \sup_{y \in S} |x - y|.$$

$\operatorname{rad}(S)$ is the radius of the smallest sphere that can fully contain S.

The *inner radius* of S is

$$r(S) = \sup_{x \in \operatorname{con} S} \inf_{\substack{T \subset S \\ T \text{ spans } x}} \operatorname{rad}(T).$$

$r(S)$ is the smallest radius of a ball centered in con S so that the ball is certain to contain points of S that span its center. $r(S)$ is a measure of the size of the nonconvexities (holes) in S.

The *inner diameter* of S, $d(S)$ is simply twice the inner radius;

$$d(S) = 2r(S).$$

Note that a sphere centered at any point, y, of con S with radius $d(S)$ contains a set of points spanning y. We make use of this property to prove

LEMMA. *Let $U \subset R^N$, U compact, $v \in R^N$, $y \in$ con U. Then there is $x \in U$ such that $|x - y| \leqslant d(U)$ and $v \cdot (x - y) \leqslant 0$.*

Proof. There are $x^j \in U, j = 1, \ldots,$ (at most $N + 1$ points x^j are required) so that for some $a^j \geqslant 0$, $\Sigma a^j = 1$, $\Sigma a^j x^j = y$ and $|x^j - y| \leqslant d(U)$. We then have

$$\Sigma a^j (x^j - y) = 0,$$

$$v \cdot \Sigma a^j (x^j - y) = 0,$$

$$\Sigma a^j v \cdot (x^j - y) = 0,$$

so for some j, $v \cdot (x^j - y) \leqslant 0$. Let this x^j be the required x. Q.E.D.

Let S^i be a countable collection of compact subsets of R^N. Further, let $d(S^i) \leqslant D$ for all i. Describe y as $\sum_{i=1}^{m} y^i$, where $y^i \in$ con S^i. Choose x^1 in S^1 so that $|y^1 - x^1| \leqslant D$. Find $x^2 \in S^2$ so that $(y^1 - x^1) \cdot (y^2 - x^2) \leqslant 0$ and $|y^2 - x^2| \leqslant D$. The lemma assures us that there is such x^2. Since $(y^2 - x^2)$ meets $(y^1 - x^1)$ at an acute angle, $|\sum_{i=1}^{2} y^i - \sum_{i=1}^{2} x^i|^2 \leqslant \sum_{i=1}^{2} |y^i - x^i|^2$. This is argued more completely in the proof of the Proposition. Hence $|\sum_{i=1}^{2} y^i - \sum_{i=1}^{2} x^i| \leqslant D(2)^{1/2}$. Proceeding sequentially, x^k is chosen in S^k within a distance D of y^k and at an acute angle to $(\sum_{i=1}^{k-1} y^i - \sum_{i=1}^{k-1} x^i)$. The lemma assures that this is possible. This gives $|\sum_{i=1}^{k} y^i - \sum_{i=1}^{k} x^i| \leqslant D(k)^{1/2}$.

Setting $k = m$, $|\sum_{i=1}^{m} y^i - \sum_{i=1}^{m} x^i| \leqslant D(m)^{1/2}$. $\sum_{i=1}^{m} x^i$ is the desired point in $\sum_{i=1}^{m} S^i$ relatively near to y. Further, as the number of sets summed, m, increases, the average discrepancy of the approximation $D(m)^{1/2}/m$ converges to 0 as $1/(m)^{1/2}$.

PROPOSITION. *Let $x \in \operatorname{con} \sum_{i=1}^{m} S^i$. Then there are $x^i \in S^i$, so that*

$$\left| x - \sum_{i=1}^{m} x^i \right| \leqslant D(m)^{1/2}.$$

Proof. The proof is by induction on m. The lemma is trivially true for $m = 1$. Suppose it is true for $m - 1$, we must demonstrate it for m.

$x = u + w$ for some u, w so that $u \in \operatorname{con} \sum_{i=1}^{m-1} S^i$, $w \in \operatorname{con} S^m$. By the inductive hypothesis there are $x^i \in S^i$, $i = 1,\dots, m-1$ so that $|u - \sum_{i=1}^{m-1} x^i| \leqslant D(m-1)^{1/2}$. Then by the lemma there is $x^m \in S^m$ so that $|w - x^m| \leqslant d(S^m) \leqslant D$ and $(x^m - w) \cdot (\sum_{i=1}^{m-1} x^i - u) \leqslant 0$. We then have

$$\begin{aligned}
&\left| \sum_{i=1}^{m} x^i - (u + w) \right|^2 \\
&\quad = \left(\sum_{i=1}^{m} x^i - (u + w) \right) \cdot \left(\sum_{i=1}^{m} x^i - (u + w) \right) \\
&\quad = \left[\left(\sum_{i=1}^{m-1} x^i - u \right) + (x^m - w) \right] \cdot \left[\left(\sum_{i=1}^{m-1} x^i - u \right) + (x^m - w) \right] \\
&\quad = \left| \sum_{i=1}^{m-1} x^i - u \right|^2 + |x^m - w|^2 + (x^m - w) \cdot \left(\sum_{i=1}^{m-1} x^i - u \right) \\
&\quad \leqslant \left| \sum_{i=1}^{m-1} x^i - u \right|^2 + |x^m - w|^2 \leqslant D^2 m. \qquad \text{Q.E.D.}
\end{aligned}$$

COROLLARY. $r(\sum_{i=1}^{m} S^i) \leqslant D(m)^{1/2}$.

The Corollary's bound is not so tight as that of the Shapley–Folkman theorem and corollaries, where it is shown that $r(\sum_{i=1}^{m} S^i) \leqslant (1/2)\, D(\min(m, N))^{1/2}$. It is derived, however, from a peculiarly simple analytic basis and proof. The Proposition and Corollary are similar to Theorem 1 of [4]. The use of statistical independence there corresponds to the possible orthogonality of $(x^m - w)$ to $(\sum_{i=1}^{m-1} x^i - u)$ above. The sequential structure of the analysis here may make this approach particularly suitable for computational use.

ACKNOWLEDGMENTS

The assistance of a John Simon Guggenheim Fellowship, NSF Grant SOC78-15429, and of the Group for the Applications of Mathematics and Statistics to Economics, University of California, Berkeley, is gratefully acknowledged. Comments of G. Debreu, E. Glustoff, and of an anonymous referee have been most helpful. They bear no responsibility for oversights.

REFERENCES

1. K. J. ARROW AND F. H. HAHN, "General Competitive Analysis," Holden–Day, San Francisco, 1972.
2. Z. ARTSTEIN, "Discrete and Continuous Bang-Bang and Facial Spaces, or: Look for the Extreme Points," Working Paper, Weizman Institute, Department of Mathematics, 1979; *SIAM Rev.*, in press.
3. J.-P. AUBIN AND I. EKELAND, "A Discrete Approach to the Bang-Bang Principle," Université de Paris, 9 Dauphine.
4. J. W. S. CASSELS, Measures of the non-convexity of sets and the Shapley–Folkman–Starr theorem, *Math. Proc. Cambridge Philos. Soc.* **78** (1975), 433–436.
5. J. GREEN AND W. P. HELLER, Mathematical analysis and convexity with applications to economics, *in* "Handbook of Mathematical Economics," North–Holland, Amsterdam, in press.
6. W. P. HELLER, Transactions with set-up costs, *J. Econ. Theory* **4** (1972), 465–478.
7. ANDREU MAS COLELL, A note on the core equivalence theorem: How many blocking coalitions are there?, *J. Math. Econ.* **5** (December 1978)December 1978), 207–216.
8. R. M. STARR, Quasi-equilibria in markets with non-convex preferences, *Econometrica* **37** (1969), 25–38.

JOURNAL OF ECONOMIC THEORY 25, 318–319 (1981)

A Note on a Comparative Statics Theorem for Choice under Risk

ELIAKIM KATZ

Department of Economics, Bar Ilan University, Ramat Gan, Israel

Received April 14, 1980; *revised* August 1, 1980

In an interesting recent paper Kraus (1979) provided a Theorem on the effect of introducing mean preserving uncertainty into θ on the optimal control α in the problem $\max_\alpha Eu(Z(\alpha, \theta))$. It is the purpose of this note to provide a simpler, shorter proof of Kraus's result as well as to somewhat generalize it. This, it is hoped, will make the result and its applications more accessible to the general economist.

THEOREM. *If* α^* *is the solution to* $\max_\alpha Eu(Z(\alpha, \theta))$, *where* θ *is a random variable with mean* $\bar{\theta}$ *and* $u' > 0$, $u'' < 0$, $Z_{\alpha\alpha} < 0$, *and if* α^{**} *is the solution to* $\max_\alpha Z(\alpha, \bar{\theta})$, *then the sign of* $\alpha^* - \alpha^{**}$ (*i.e., the impact of uncertainty on the optimal* α) *is given by the following rule*:

$Z_{\alpha\theta\theta}$	$Z_{\alpha\theta} \cdot Z_\theta$	$\operatorname{sign}(\alpha^* - \alpha^{**})$
+	+	?
+	−	+
+	0	+
−	+	−
−	−	?
−	0	−
0	+	−
0	−	+
0	0	0

Proof. For our proof it is simply necessary to determine the sign of $A = Z_\alpha(\alpha^*, \bar{\theta})$. If A is positive (negative) (zero), then by $Z_{\alpha\alpha} < 0$ this implies that α^* is smaller than (larger than) (equal to) α^{**}.

Now, since $Eu' > 0$ and since A is non-random, the sign of A is as the sign of $B = A \cdot Eu'(Z(\alpha^*, \theta)) = EA \cdot u'(Z(\alpha^*, \theta))$. Using a Taylor Expansion, A

0022-0531/81/050318-02$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

may be written (exactly) as $Z_\alpha(\alpha^*, \theta) + (\bar{\theta} - \theta) \cdot Z_{\alpha\theta}(\alpha^*, \theta + \gamma(\bar{\theta} - \theta))$, where $0 < \gamma < 1$. Hence, B may be written as

$$B = \underset{(C)}{E(Z_\alpha(\alpha^*, \theta) \cdot u'(Z(\alpha^*, \theta)))} + \underset{(D)}{E((\bar{\theta} - \theta)H)},$$

where $H = Z_{\alpha\theta}(\alpha^*, \theta + \gamma(\bar{\theta} - \theta)) \cdot u'(Z(\alpha^*, \theta))$.

By the first order conditions $(C) = 0$. Furthermore, $\bar{\theta} - \theta$ is a random variable with mean zero. Hence, $B = E((\bar{\theta} - \theta)H) =$ covariance $(\bar{\theta} - \theta, H)$. This will be positive (negative) if $\partial(\bar{\theta} - \theta)/\partial\theta = -1$, and $\partial H/\partial\theta = u' \cdot (1 - \gamma) \cdot Z_{\alpha\theta\theta}(\alpha^*, \theta + \gamma(\bar{\theta} - \theta)) + u'' \cdot Z_\theta(\alpha^*, \theta) \cdot Z_{\alpha\theta}(\alpha^*, \theta + \gamma(\bar{\theta} - \theta))$ have the same (opposite) sign and will be zero if $\partial H/\partial\theta = 0$. Hence, $\alpha^* - \alpha^{**}$ has the same sign as $(\partial H/\partial\theta)$ above. The Theorem immediately follows.

REFERENCES

M. KRAUS (1979), A comparative statics theorem for choice under risk, *J. Econ. Theory* **21**, 510–517.

JOURNAL OF ECONOMIC THEORY 25, 320–321 (1981)

Papers to Appear in Forthcoming Issues

The following articles are among those newly accepted for publication in the Journal of Economic Theory. They are listed alphabetically by first author.

L. E. BLUME, M. M. BRAY, AND D. EASLEY, Introduction to the Stability of Rational Expectations Equilibrium. Correspondent: Professor David Easley, Department of Economics, Cornell University, Ithaca, NY 14853

LAWRENCE E. BLUME AND DAVID EASLEY, Learning to be Rational. Correspondent: Professor David Easley, Department of Economics, Cornell University, Uris Hall, Ithaca, NY 14853.

LAWRENCE BLUME, DAVID EASLEY, AND MAUREEN O'HARA, Characterization of Optimal Plans for Stochastic Dynamic Programs. Correspondent: Dr. David Easley, Department of Economics, Cornell University, Uris Hall, Ithaca, NY 14853.

MARGARET BRAY, Learning, Estimation, and the Stability of Rational Expectations. Correspondent: Professor Margaret Bray, Faculty of Economics and Politics, Cambridge University, Sidgwick Avenue, Cambridge, England, CB3 9DD.

PHILIP H. DYBVIG, Recovering Preferences from Preferences over Nominal Gambles. Correspondent: Professor Philip H. Dybvig, School of Organization and Management, Yale University, Box 1A, New Haven, CT 06520.

DUNCAN K. FOLEY, Realization and Accumulation in a Marxian Model of the Circuit of Capital. Correspondent: Professor Duncan K. Foley, Department of Economics, Barnard College, New York, NY 10025.

MARTIN F. HELLWIG, Rational Expectations Equilibrium with Conditioning on Past Prices: A Mean Variance Example. Correspondent: Professor Martin F. Hellwig, Wirtschaftstheoretische Abteilung III, Universität Bonn, Adenauerallee 24, D-53, Bonn, West Germany.

JAMES S. JORDAN, The Competitive Allocation Process is Informationally Efficient Uniquely. Correspondent: Professor J. S. Jordan, Department of Economics, University of Minnesota, 271 19th Avenue South, Minneapolis, MN 55455.

JAMES S. JORDAN, Admissible Market Data Structures: A Complete Characterization. Correspondent: Professor James S. Jordan, Department of Economics, University of Minnesota, Minneapolis, MN 55455.

JAMES S. JORDAN, A Dynamic Model of Expectations Equilibrium. Correspondent: Professor J. S. Jordan, Department of Economics, University of Minnesota, Minneapolis, MN 55455.

JAMES S. JORDAN, The Generic Existence of Rational Expectations Equilibrium in the Higher Dimension Case. Correspondent: Professor J. S. Jordan, Department of Economics, University of Minnesota, Minneapolis, MN 55455.

M. ALI KHAN AND SALIM RASHID, Approximate Equilibria in Markets with Indivisible Commodities. Correspondent: Professor M. Ali Khan, Department of Political Economy, The Johns Hopkins University, Baltimore, MD 21218.

MORDECAI KURZ AND SERGIU HART, Pareto Optimal Nash Equilibria are Competitive in a Repeated Economy. Correspondent: Professor Mordecai Kurz, IMSSS, Stanford University, Fourth Floor, Encina Hall, Stanford, CA 94305.

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

JOHN LAITNER, The Definition of Stability in Models with Perfect Foresight. Correspondent: Professor John Laitner, Department of Economics, University of Michigan, Ann Arbor, MI 48109.

DAVID C. NACHMAN, Preservation of "More Risk Averse" under Expectations. Correspondent: Professor David C. Nachman, College of Management, Georgia Institute of Technology, Atlanta, GA 30332.

YOSHIHIKO OTANI, A Simple Proof of the LeChatelier–Samuelson Principle and the Theory of Cost and Production. Correspondent: Professor Yoshihiko Otani, Department of Economics, University of Kansas, Lawrence, KS 66045.

DAN PELED, Informational Diversity over Time and the Optimality of Monetary Equilibria. Correspondent: Professor Dan Peled, Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA 15213.

RAFAEL ROB, Asymptotic Efficiency of the Demand Revealing Mechanism. Correspondent: Professor Rafael Rob, Department of Economics, Brown University, Providence, RI 02912.

GARRY J. SCHINASI, Fluctuations in a Dynamic, Intermediate-Run IS-LM Model: Applications of the Poincaré Bendixon Theorem. Correspondent: Dr. Garry J. Schinasi, Board of Governors of the Federal Reserve System, Room B-3245 C, Washington, D.C. 20551.

DAVID SCHMEIDLER, A Condition Guaranteeing that the Nash Allocation is Walrasian. Correspondent: Professor David Schmeidler, P.O. Box 39286, Tel Aviv, Israel, 61392.

WILLIAM THOMSON, A Class of Solutions to Bargaining Problems. Correspondent: Professor William Thomson, Department of Economics, Harvard University, Cambridge, MA 02138.

PAUL A. WELLER, The Speed of Convergence of Prices in Random Exchange Economies. Correspondent: Professor Paul A. Weller, Department of Economics, Warwick University, Coventry, England, CV4 7AL.

Printed in Belgium

ACADEMIC PRESS

Introducing the first principles text with a supply-side emphasis—

INTRODUCTORY ECONOMICS

by MICHAEL VESETH, University of Puget Sound

January 1981

THE COMPLETE PACKAGE paperbound Macro/Micro volumes. Coursebook. Instructor's Manual. Test Book. Computerized Test File. Transparency Masters.

PRAISE FOR VESETH'S INTRODUCTORY MACROECONOMICS

"It satisfies the main thing I am looking for: a readable text that doesn't talk down to students and that approaches macroeconomics from an aggregate supply-aggregate demand angle instead of the antiquated income-expenditure approach. . . ."—Professor Kenneth L. Cavanaugh, U. of Washington, Seattle.

ECONOMICS: Private and Public Choice

Second Edition

JAMES D. GWARTNEY, Florida State University, and
RICHARD STROUP, Montana State University
with the assistance of A. H. Studenmund, Occidental College.

1980 *835 pages*

"Excellent book—perhaps the best available. You have kept the features that led me to adopt the first edition: style, organization, and topics—particularly the unique coverage of public choice literature as an extension of the 'economic way of thinking.' The coursebook is the best available. . . ."

The Supplements: two paperbound volumes: MACROECONOMICS: Private and Public Choice, 2e; MICROECONOMICS: Private and Public Choice, 2e/ Coursebook/ Test Bank/Computerized Test File/Instructor's Manual/Transparency Masters.

MONEY, BANKING, AND ECONOMIC ACTIVITY

by GAIL E MAKINEN, Principal Macroeconomist,
U.S General Accounting Office, and Professorial Lecturer,
Georgetown University.

January 1981

An up-to-date look at money and banking, incorporating the 1980 finance laws.

For examination copies, write to College Department, Academic Press, Inc , 111 Fifth Avenue, New York, N Y 10003 Please state course title, enrollment, and present text

ACADEMIC PRESS, INC.
A Subsidiary of Harcourt Brace Jovanovich, Publishers
New York • London • Toronto • Sydney • San Francisco

AP 7965

FORECASTING IN BUSINESS AND ECONOMICS

By C. W. J. GRANGER

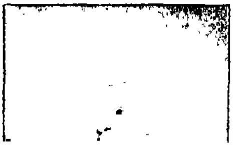
University of California, San Diego

Written by the recognized authority on forecasting theory and application, *Forecasting in Business and Economics* gives the senior, MBA, and beginning graduate student a clear cut strategy for tackling forecasting problems

Unlike any other book in the field, this new text stresses the importance of selecting a relevant information set and methods of evaluating the forecast once it has been prepared Granger's emphasis on critical assessment gives the future economist and manager a sound base for interpreting forecasts and making decisions about the future

Modern techniques of statistical, econometric, population, and technological forecasting are presented along with discussions of a wide range of topics including Box-Jenkins methods, leading indicators, survey data, world models and the use of information sets of increasing size

1980 226 pages



FUNDAMENTALS OF BUSINESS STATISTICS

JEROME D BRAVERMAN, Rider College

A business statistics book that the job-oriented student can really profit from! For each statistical technique introduced, an example of its application is given with a step-by step solution Students gain a true perspective of how statistical techniques are used in the business world while they gain practice in intelligent decision making

A special *review chapter* early in the book gives students with weak backgrounds the confidence they'll need to succeed in the course Only a knowledge of high school math is required

A *Student Solutions Manual* is available

1979 597 pages

ACADEMIC PRESS, INC.

A Subsidiary of Harcourt Brace Jovanovich Publishers

NEW YORK LONDON TORONTO SYDNEY SAN FRANCISCO

For examination copies write to College Department ACADEMIC PRESS INC 111 Fifth Avenue, New York, N Y 10003 Please state course title enrollment and present text

AP 7964

How To Comply With The New Copyright Law

Libraries everywhere have found the easy way to fill photocopy requests legally and instantly, without the need to seek permissions, from this and over 3000 other key publications in business, science, humanities, and social science

Participation in the Copyright Clearance Center (CCC) assures you of legal photocopying at the moment of need You can

Fill requests for multiple copies, interlibrary loan (beyond the CONTU guidelines), and reserve desk without fear of copyright infringement

Supply copies simply and easily from registered publications The CCC's flexible reporting system accepts photocopying reports and returns an itemized invoice. *You need not keep any records*, our computer will do it for you

The Copyright Clearance Center is your one-stop place for on-the-spot clearance to photocopy for internal use You will never have to decline a photocopy request or wonder about compliance with the law for any publication registered with the CCC

For more information, just contact

Copyright Clearance Center
21 Congress Street
Salem, Massachusetts 01970
(617) 744-3350
a not-for-profit corporation

NAME TITLE

ORGANIZATION

ADDRESS

CITY STATE ZIP

COUNTRY TELEPHONE

JOURNAL OF ECONOMIC THEORY

Volume 25, Number 3, December 1981

Copyright © 1981 by Academic Press, Inc.
All Rights Reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc. (21 Congress Street, Salem, Massachusetts 01970), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre-1981 articles are the same as those shown for current articles.

Published bimonthly at 41 Tempelhof, Bruges, Belgium by
Academic Press, Inc., 111 Fifth Avenue, New York, N.Y. 10003.

1981: Volumes 24–25. Price $138.00 U.S.A.; $165.00 outside U.S.A.
1982: Volumes 26–28. Price: $192.00 U.S.A.; $229.50 outside U.S.A.
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publisher. All correspondence and subscription orders should be addressed to the office of the Publishers at 111 Fifth Avenue, New York, N.Y. 10003. Send notices of change of address to the office of the Publishers at least 6 to 8 weeks in advance. Please include both old and new addresses. Postmaster: Send changes of address to *Journal of Economic Theory*, 111 Fifth Avenue, New York, New York 10003.

Second class postage paid at Jamaica, N.Y. Air freight and mailing in the U.S.A. by Publications Expediting, Inc., 200 Meacham Avenue, Elmont, New York 11003.

Copyright © 1981 by Academic Press, Inc.
Printed in Bruges, Belgium, by the St. Catherine Press, Ltd.

JOURNAL OF ECONOMIC THEORY **25**, ii (1981)

Notice

We are pleased to announce that in 1982, *Journal of Economic Theory* will be expanded to three volumes of two issues each. During 1982, Volumes 26 through 28 will be published; the frequency of publication will be bimonthly with the first issue of the year appearing in February.

Copyright © 1981 by Academic Press, Inc.

JOURNAL OF ECONOMIC THEORY 25, 323-337 (1981)

Constrained Excess Demand Functions*

HERAKLIS M. POLEMARCHAKIS†

Department of Economics, Columbia University, New York, New York 10027,
and C.O.R.E., Université Catholique de Louvain, Louvain la Neuve, Belgium

Received January 5, 1979; revised May 23, 1980

1. INTRODUCTION

It is the starting point of non-Walrasian equilibrium theory that agents perceive constraints on the amount they can transact in the various markets. This formulation follows naturally from the observation that, away from a Walrasian equilibrium price system, agents' Walrasian—notional—excess demands are incompatible, and hence, if economic activity is to occur, some agents must be rationed out of their desired level of transaction. The assumption of rationality of individual behavior then implies that agents will express excess demand for the various commodities taking into consideration the prevailing rationing mechanism.

The question I take up in this paper concerns the restrictions that are imposed on individual behavior under quantity constraints by the assumption of rationality:[1] Consider an agent who expresses constrained—effective—excess demand so as to maximize a monotone, strictly quasi-concave preference pre-order, subject to binding quantity constraints in a subset K of the system of markets, as well as a budget constraint. The constrained excess demand vector x^K is a function of the price vector p, and the vector of quantity constraints s, $x^K(p, s)$. Furthermore, the vector x^K can be decomposed into a vector $x^K_{K'}$, of the constrained excess demand for goods in the markets for which the agent is not constrained (K' is the subset complementary to K) and a vector x^K_K of the constrained excess demand for goods in the markets for which the agent is indeed constrained. The function $x^K_K(p, s)$ is trivial, it is the projection $x^K_K(p, s) = s$. What needs to be characterized is the constrained excess demand function in the unconstrained

---

* This work was supported by National Science Foundation Grant SOC77-06000 at the Institute for Mathematical Studies in the Social Sciences, Stanford University.

† I would like to thank Marie Blanchard for insightful comments. I am grateful to Kenneth Arrow, Jean Jacques Laffont and Andreu Mas-Colell for very helpful discussions.

[1] In a different framework, the same problem has been taken up by Diamond and Yaari (1972) and Tobin (1952).

0022-0531/81/060323-15$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

markets—$x^K_{K'}(p, s)$. In particular, one can raise the question whether an arbitrary function can be generated as the constrained excess demand function, in the unconstrained markets, of a rational agent. If the problem is approached in an "infinitesimal" framework, the question of characterization of the function $x^K_{K'}(p, s)$ is reduced to a characterization of the Jacobian matrices $D_p x^K_{K'}(p, s)$ and $D_s x^K_{K'}(p, s)$. The characterization of $D_p x^K_{K'}(p, s)$ is an immediate extension of the characterization of the Jacobian of a notional excess demand function through the decomposition into a substitution matrix and a matrix of income effects. The matrix $D_s x^K_{K'}(p, s)$ on the other hand, has no counterpart in Walrasian theory. It is usually referred to as the *spill-over matrix* and captures the change in the excess demand of an agent as the quantity constraints change. I shall demonstrate that no restrictions (other than adding up, of course) can be imposed on the spill-over matrix;[2] as long as one wants to consider the entire class of monotone strictly quasi-concave preference pre-orders, the spill-over matrix must remain arbitrary.[3]

It is a widely accepted—even if little understood—generalization that the immediate impact of fluctuations in supply and demand is reflected in quantity, rather than price, adjustments. As a consequence, qualitative analysis of short-run dynamics, as well as short-run policy recommendations depend on the properties of the spill-over matrix.[4] Furthermore, restrictions on the spill-over coefficients are necessary for the estimation of multimarket disequilibrium models.[5] In this framework, it is rather unpleasant that consumer theory fails to provide the required structure. It is an open problem to determine plausible restrictions on individual characteristics that yield the desired behavior under quantity constraints. These restrictions must involve statements concerning the nature of the dependence of the marginal utility of goods h, $h \in K'$, on the level of consumption of goods h', $h' \in K$. In other words, they must characterize the off-diagonal elements of the Hessian matrix of the utility function. It is the fact that concavity imposes no restrictions on the off-diagonal elements that leads to the results derived in the present paper.

Remark. That rationality imposes no restrictions on the spill-over matrix for a particular state of the markets (i.e., a particular K and s) does *not*, of course, imply that the spill-over matrices corresponding to different states of the markets can be independently specified. Similarly, the indeterminacy of the spill-over matrix for the definition of effective demand adopted here does *not* imply that alternative formulations (e.g., Benassy (1975), Clower (1965), Laroque (1976, 1978), Malinvaud (1977)) display indeterminacy as well.

---

[2] Observe that this result does *not* depend on aggregation. It is in the case of notional excess demands that aggregation leads to arbitrary behavior—see Geanakoplos and Polemarchakis (1977) and Mantel (1977) for surveys of this problem.

[3] This result does *not* affect the argument in Laroque (1976, 1978) or Wiesmeth (1977). That the aggregate spill-over matrix has generically fully rank is clearly compatible with the result derived here.

[4] See Malinvaud (1977).

[5] See Ito (1978).

2. CONSTRAINED CONSUMER EXCESS DEMAND FUNCTIONS

Consider an economy with $(l+1)$ commodities indexed by h, $h = 0, 1,..., l$. A consumer is characterized by a consumption set X, a convex subset of $\mathscr{R}^{l+1}$, a strictly quasi-concave utility function u defined on X, and a vector of initial endowments w in X. A price system is a vector p in $\mathscr{R}^{l+1}_+$. In addition to the price system, a trader perceives for each commodity h, other than 0, quantitative constraints, $\bar{z}_h \geqslant 0$, $\underline{z}_h \leqslant 0$, that set upper and lower bounds on the amount of commodity h he can trade. It is clear that the consumer must take the quantity constraints into account in expressing his (effective) excess demand; yet, no general agreement has been reached on the price formulation of the consumer's decision problem. In particular, it is not clear whether a trader does (or should) take into consideration the constraints he faces in the market for good k when expressing his excess demand for k. I shall follow here the formulation suggested by Drèze.[6] A trader (X, u, w) expresses excess demand x by solving the problem

$$\begin{aligned} &\operatorname*{Max}_{x \in X - \{w\}} \; u(w+u) \\ &\text{s.t. } p'x = 0, \\ &\qquad \underline{z}_h \leqslant x_h \leqslant \bar{z}_h, \qquad h = 1,..., l.^{7} \end{aligned}$$

He is said to be constrained on market k if there exists a consumption bundle $(w + x')$ in X satisfying $p'x' = 0$ and the quantitative constraint $\underline{z}_h \leqslant x'_h \leqslant \bar{z}_h$ for all h different from 0 and k, which provides a higher utility than x. If $x'_k > \bar{z}_k$ the trader is said to be constrained in the demand for k; if $x'_k < \underline{z}_k$, he is said to be constrained in the supply of k. Let K be the set of markets on which the trader is constrained, and let $s \in \mathscr{R}^{|K|}$ be the vector of quantity constraints—for $h \in K$, s_h is equal to $\bar{z}_h$ or $\underline{z}_h$ depending on whether the trader is constrained in the demand or the supply, respectively, of commodity h. The constrained excess demand x is then the vector in $X - \{w\}$ which maximizes $u(x + w)$ subject to the budget constraint $p'x = 0$ and the quantity constraints $x_h = s_h$ for $h \in K$. Observe that I am now writing the quantity constraints with equality. This allows the constrained excess demand function $x^K(p, s)$ to be differentiable with respect to the vector of constraints, s. On the other hand, $x^K(p, s)$ is not necessarily defined for an arbitrary s; furthermore, maximization under the constraint $|x_h| \leqslant |s_h|$ is not necessarily equivalent to maximization under the constraint $x_h = s_h$.[8] These problems need not be explicity dealt with here, however, since I shall only be interested at the infinitesimal characterization of the constrained excess demand function.

---

[6] See Drèze (1975).

[7] For an alternative formulation see Benassy (1975) and Clower (1965).

I shall make the following assumptions concerning the consumer's characteristics:

ASSUMPTION 1. $X = \mathscr{R}^{l+1}_{+}$.

ASSUMPTION 2. *The utility function* u *is a twice continuously differentiable and strictly quasi concave function from* X *to* $\mathscr{R}$. *At all* $x \in X$, $Du(x) \gg 0$, *the matrix* $D^2u(x)$ *is negative definite on the orthogonal complement of* $Du(x)$, $[Du(x)]^{\perp}$, *and the closure of the indifference hypersurface through* x *is contained in* X.

Remark. Since $D^2u(x)$ is negative definite on $[Du(x)]^{\perp}$, for any $K' \subset \{0, 1, \ldots, l\}$ the matrix $D^2_{K'K'}u(x)$ is negative definite on $[D_{K'}u(x)]^{\perp}$ and the matrix $\left|\begin{smallmatrix} D^2_{K'K'} & D_{K'}u(x) \\ D_{K'}u(x)^t & 0 \end{smallmatrix}\right|$ if of full rank.

In the discussion that follows I hold the consumption set fixed and, consequently, I refer to a trader as an ordered pair (u, w).

Consider the decision problem of an agent (u, w) who perceives quantity constraints in the markets for commodities indexed by h, for $h \in K$. K is a subset of index set $\{1, \ldots, l\}$, and K' is the set $\{0, 1, \ldots, l\}/K$. Let $p = (p_0, p_1, \ldots, p_l) \in \mathscr{R}^{l+1}_{+}$ denote the price vector and $s = (s_h \mid h \in K)$ denote the vector of quantity constraints faced by the agent. The agent then solves the problem

$$\begin{aligned} &\underset{x \in X - \{w\}}{\text{Max}} \quad u(x + w) \\ &\text{s.t. } x_h = s_h, \qquad h \in K, \qquad\qquad (1) \\ &\quad\;\; p^t x = 0. \end{aligned}$$

By Assumption 2, given $p \gg 0$ and s such that $(w_h + s_h) > 0$ for all $h \in K$ and $p^t(y - w) < 0$ for some $y \in \mathscr{R}^{l+1}_{+}$ with $y_h = (s_h + w_h)$ for $h \in K$, a solution to (1) exists and is unique. Furthermore, by the Kuhn–Tucker theorem, x^* solves (1) if and only if there exist real numbers λ^*, μ^*_h, $h \in K$ such that

---

[8] See Laroque ((1976, pp. 10–12)) for a discussion.

$$
\begin{aligned}
D_h u(x^* + w) - \lambda^* p_h &= 0, & h &\in K',\\
D_h u(x^* + w) - \lambda^* p_h + \mu_h^* &= 0, & h &\in K,\\
x_h^* &= s_h, & h &\in K,\\
p' x^* &= 0.
\end{aligned}
\tag{2}
$$

Totally differentiating the system of equations (2) we get

$$
\begin{bmatrix}
D^2 u & -p & e^{h_1+1} & \cdots & e^{h_{|K|}+1}\\
-p' & 0 & 0 & \cdots & 0\\
(e^{h_1+1})' & 0 & 0 & \cdots & 0\\
\vdots & \vdots & \vdots & & \vdots\\
(e^{h_{|K|}+1})' & 0 & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
dx^*\\ d\lambda^*\\ d\mu_{h_1}^*\\ \vdots\\ d\mu_{h_{|K|}}^*
\end{bmatrix}
=
\begin{bmatrix}
\lambda^* dp\\ (dp)' x^*\\ ds_{h_1}\\ \vdots\\ ds_{h_{|K|}}
\end{bmatrix},
\tag{3}
$$

where $\{h_1, \ldots, h_{|K|}\} = K$ in the natural order, and e^i is the ith unit vector in $\mathscr{R}^{l+1}$. Without loss of generality we may assume that $K = \{k, \ldots, l\}$ for some $k \in \{1, \ldots, l\}$. Furthermore, letting

$$
D^2 u = \begin{bmatrix} D_{K'K'} & D_{K'K}\\ D_{KK'} & D_{KK} \end{bmatrix} \quad \text{and} \quad p = \begin{bmatrix} p_{K'}\\ p_K \end{bmatrix}
$$

we can rewrite (3) as

$$
\begin{bmatrix}
D^2_{K'K'} u & D^2_{K'K} u & -p_{K'} & 0\\
D^2_{KK'} u & D^2_{KK} u & -p_K & I\\
-p'_{K'} & -p'_K & 0 & 0\\
0 & I & 0 & 0
\end{bmatrix}
\begin{bmatrix}
dx^*_{K'}\\ dx^*_K\\ d\lambda^*\\ d\mu^*
\end{bmatrix}
=
\begin{bmatrix}
\lambda^* dp_{K'}\\ \lambda^* dp_K\\ (dp)' x^*\\ ds
\end{bmatrix}.
\tag{3'}
$$

It is easy to see, by elementary row and column operations, that the matrix

$$
A = \begin{bmatrix}
D^2_{K'K'} u & D^2_{K'K} u & -p_{K'} & 0\\
D^2_{KK'} u & D^2_{KK} u & -p_K & I\\
-p'_{K'} & -p'_K & 0 & 0\\
0 & I & 0 & 0
\end{bmatrix}
$$

is similar to the matrix

$$
B = \begin{bmatrix}
D^2_{K'K'} u & -p_{K'} & 0\\
-p'_{K'} & 0 & 0\\
0 & 0 & I
\end{bmatrix}.
$$

Consequently, the matrix A is invertible if and only if the matrix B is invertible. Furthermore, the matrix B is invertible if and only if the matrix

$$C = \begin{bmatrix} D^2_{K'K'}u & -p_{K'} \\ -p'_{K'} & 0 \end{bmatrix}$$

if of full rank. But, from (2), $\lambda^* p_{K'} = D_{K'}u$. Since, by Assumption (2), $Du \gg 0$ everywhere on $\dot{\mathscr{R}}^{l+1}_{+}$, $\lambda^* > 0$, and hence,

$$\det|C| = \left(\frac{1}{\lambda^*}\right)^2 \det[D],$$

where

$$D = \begin{bmatrix} D^2_{K'K'}u & D_{K'}u \\ D_{K'}u & 0 \end{bmatrix}.$$

But, by definition, $\det|D|$ is the Gaussian curvature of the hypersurface through $(w + x^*)$ of the restriction of u to the coordinate subspace $T(K, w, s) = \{y \in \dot{\mathscr{R}}^{l+1}_{+} \mid y_h = w_h + s_h \text{ for all } h \in K\}$. By Assumption (2) then, the matrix D has non-zero determinant, and hence is invertible, which implies the invertibility of the matrix A.[9]

Consider now the solution to (1) as a function of p and s, $x^K(p, s)$. It is a well-defined and continuous function from $S(K, w)$ to $\mathscr{R}^{l+1}$, where $S(K, w) = \{(p, s) \in \dot{\mathscr{R}}^{l+1}_{+} \times \mathscr{R}^{|K|} \mid (w_h + s_h) > 0$ for $h \in K$, and $p'(y - w) < 0$ for some $y \in \dot{\mathscr{R}}^{l+1}_{+}$ with $y_h = (s_h + w_h)$ for $h \in K\}$. Furthermore, the invertibility of the matrix A and the continuous dependence of the inverse on (p, s) are sufficient to guarantee that the constrained excess demand function is continuously differentiable. We have demonstrated the following:[10]

PROPOSITION 1. *Under Assumptions* (1) *and* (2) *the constrained excess demand function* $x^K(p, s)\colon S(K, w) \to \mathscr{R}^{l+1}$ *is continuously differentiable.*

Remark. The function $x^K(p, s)$ is, of course, homogeneous of degree 0 in p.

Remark. The functions $\lambda^K(p, s)$ and $\mu^K_h(p, s)$, $h \in K$, are, similarly, continuously differentiable; they are homogeneous of degree (-1) in p.

The discussion has, up to this point, been limited to the formulation of the agent's constrained maximization problem and the demonstration of the fact that the regularity properties imposed on individuals preferences by Assumption (2) are sufficient to guarantee the differentiability of the constrained excess demand function $x^K(p, s)$. The question follows whether the regularity assumptions have any further qualitative implications concerning the agent's response to a change either in prices or in the quantity constraints it perceives.

---

[9] Given a matrix A with rows $i = 1, \ldots, l_1$ and columns $j = 1, \ldots, l_2$ and subsets $K_1 \subset \{1, \ldots, l_1\}$, $K_2 \subset \{1, \ldots, l_2\}$, A_{K_1, K_2} denotes the obvious submatrix. Similarly, given a vector $y = (y_1, \ldots, y_l)$ and $K \subset \{1, \ldots, l\}$, y_k denotes the obvious vector in $\mathscr{R}^{|K|}$.

[10] This is, of course, an immediate extension of the argument in Debreu (1972).

The Jacobian of the function $x^K(p, s)$ is determined by the matrix A^{-1}. Since the matrix A is symmetric, A^{-1} is symmetric as well, and hence with no loss of generality we may write

$$A^{-1} = \begin{bmatrix} S_{K'K'} & S_{K'K} & -v_{K'} & R_{K'} \\ (S_{K'K})^t & S_{KK} & -v_K & R_K \\ (-v_{K'})^t & (-v_K)^t & b_{00} & -b_K^t \\ (R_{K'})^t & (R_K)^t & -b_K & B_K \end{bmatrix}.$$

Since the prouct $AA^{-1} = I$, the identity matrix, the following conditions must be satisfied:

(a) $S_{K'K} = 0; S_{KK} = 0.$
(b) $v_K = 0.$
(c) $p_{K'}^t(S_{K'K'}) = (S_{K'K'})p_{K'} = 0.$
(d) $p_{K'}^t v_{K'} = 1.$
(e) $R_K = I.$
(f) $p_{K'}^t R_{K'} = -p_K^t.$

Furthermore, the following argument shows that the matrix $S_{K'K'}$ is negative semi-definite and has rank $(|K'| - 1)$. Since $A^{-1}A = I$, $(D_{K'K'}^2 u)(S_{K'K'}) + p_{K'}v_{K'}^t = I$. Let y be any non-zero eigenvector of $S_{K'K'}$ and suppose $(S_{K'K'})y = \mu y$, for some $\mu \neq 0$. Observe that y is in the span of the columns of $S_{K'K'}, [S_{K'K'}]$,and hence $y^t p_{K'} = 0$. Consequently, $y^t[(D_{K'K'}^2 u)(S_{K'K'})]y = y^t y$, and hence $\mu y^t(D_{K'K'}^2 u)y = y^t y$. Since the matrix $D^2 u$ is negative definite on the orthogonal complement of Du, the matrix $D_{K'K'}^2 u$ is negative definite on the orthogonal complement of $D_{K'} u$, $[D_{K'} u]^{\perp}$. But, from (2), $[D_{K'} u] = [p_{K'}]$, and, since $y^t p_{K'} = 0$, $y \in [D_{K'u}]^{\perp}$. As a result, $y^t(D_{K'K'}^2 u)y < 0$, and, since $y^t y > 0$, the eigenvalue μ of $S_{K'K'}$ must be negative. It remains to show that rank $(S_{K'K'}) = (|K'| - 1)$. But this follows from the fact that $p_{K'}^t S_{K'K'} = 0$, while for all $y \in [p_{K'}]^{\perp}$, $y^t[(D_{K'K'}^2 u)(S_{K'K'})]y = y^t y$ and $y^t(D_{K'K'}^2 u) \neq 0$. Finally, since the matrix $S_{K'K'}$ is negative semi-definite, has rank $(|K'| - 1)$ and satisfies $p_{K'}^t(S_{K'K'}) = 0$ with $p_{K'}^t \neq 0$, $S_{K'K'}$ is negative definite when restricted to the complement of $[p_{K'}]$. We have thus demonstrated the following:

Proposition 2. *Let $x^K(p, s)$ be the constrained excess demand function of an agent (u, w) satisfying Assumptions* (1) *and* (2). *Then everywhere on $S(K, w)$, $p^t x^K(p, s) = 0$, $x^K_K(p, s) = s$, and the following conditions are satisfied*:

(1) $D_{p_{K'}} x^K_{K'} = \lambda S_{K'K} - v_{K'}(x^K_{K'})^t$; *the matrix $S_{K'K'}$ is symmetric, negative semi-definite of rank $(|K'| - 1)$ and satisfies $p^t_{K'} S_{K'K'} = S_{K'K'} p^t_{K'} = 0$; the scalar λ is strictly positive; the vector $v_{K'}$ satisfies $p^t_{K'} v_{K'} = 1$*;

(2) $D_{p_K} x^K_K = 0$; $D_{p_K} x^K_{K'} = -v_{K'} x^t_K$; $D_{p_K} x^K_K = 0$;

(3) $D_s x^K_{K'} = R_{K'}$; *the matrix $R_{K'}$ satisfies $p^t_{K'} R_{K'} = -p^t_K$*;

(4) $D_s x^K_K = I$.

Remark. Homogeneity of degree zero with respect to p of the function $x^K(p, s)$ holds, if $dp = \alpha p$, $\alpha > 0$ and $ds = 0$ imply $dx^K = 0$. But, since $dx^K_{K'} = \lambda S_{K'K'} dp_{K'} - v_{K'}(x^K)^t dp + R_{K'} ds$ while $dx^K_K = ds$, and furthermore $S_{K'K'} p_{K'} = 0$ and $(x^K)^t p = 0$, this is indeed the case.

Remark. Observe that the only restriction claimed on the spill-over matrix $R_{K'}$ is the condition $p^t_{K'} R_{K'} = -p^t_K$. This restriction of course only says that the consumer stays on the budget hyperplane; it is equivalent to $\sum_{h' \in K'} p_{h'}(\partial x^K_{h'}/\partial s_h) + p_h = 0$, for all $h \in K$.

Proposition 2 establishes a set of necessary conditions for the function $x^K(p, s)$ to be the constrained excess demand function of a rational agent. The question I want to take up now is whether the set of restrictions imposed by Proposition 2 is not only necessary but also sufficient for rationality. It is clear that since conditions (1) to (4) of Proposition 2 refer only to the Jacobian of constrained excess demand function, the strongest result that can be hoped for is that they are sufficient for the function $x^K(p, s)$ to be infinitesimally the constrained excess demand function of a rational agent. The following argument shows that this is indeed the case.

Suppose we are given a vector $\bar{x}^K$ in $\mathscr{R}^{l+1}$, a vector $\bar{s}$ in $\mathscr{R}^{|K|}$, a vector $\bar{p}$ in $\mathscr{R}^{l+1}_+$, a $(|K'| \times |K'|)$ matrix $\bar{S}_{K'K'}$, a vector $\bar{v}_{K'}$, in $\mathscr{R}^{|K'|}$ and a $(|K'| \times |K|)$ matrix $\bar{R}_{K'}$, such that $\bar{x}^K_K = \bar{s}$, $\bar{p}^t \bar{x}^K = 0$ and the following conditions are satisfied:

(a) The matrix $\bar{S}_{K'K'}$ is symmetric, negative semi-definite of rank $(|K'| - 1)$ and satisfies $\bar{p}^t_{K'} \bar{S}_{K'K'} = \bar{S}_{K'K'} \bar{p}_{K'} = 0$; the vector $\bar{v}_{K'}$ satisfies $\bar{p}^t_{K'} \bar{v}_{K'} = 1$.

(b) The matrix $\bar{R}_{K'}$, satisfies $\bar{p}^t_{K'} \bar{R}_{K'} = -\bar{p}^t_K$.

Consider the matrix

$$E = \begin{bmatrix} \bar{S}_{K'K'} & 0 & -\bar{v}_{K'} & \bar{R}_{K'} \\ 0 & 0 & 0 & I \\ -\bar{v}^t_{K'} & 0 & b_{00} & -b^t_K \\ \bar{R}^t_{K'} & I & -b_K & B_K \end{bmatrix},$$

where the scalar b_{00}, the vector b_K and the matrix B_K are chosen in a way to be specified shortly. Since $\bar{p}_{K'}\bar{S}_{K'K'} = 0$ while $\bar{p}^t_{K'}\bar{v}_{K'} = 1$, $\bar{v}_{K'} \notin [\bar{S}_{K'K'}]$. Consequently, the matrix

$$F = \begin{bmatrix} \bar{S}_{K'K'} & -\bar{v}_{K'} \\ -\bar{v}^t_{K'} & 0 \end{bmatrix}$$

has full rank, $(|K'| + 1)$. By elementary column and row operations it is easy to see that the matrix E is similar to the matrix

$$G = \begin{bmatrix} F & 0 \\ 0 & I \end{bmatrix}$$

and, hence, is of full rank. Let E^{-1} be the inverse matrix: By symmetry, E^{-1} can be written as follows:

$$E^{-1} = \begin{bmatrix} \bar{U}''_{K'K'} & \bar{U}''_{K'K} & -q_{K'} & d^k_{K'} & \cdots & d^l_{K'} \\ (\bar{U}''_{K'K})^t & \bar{U}''_{KK} & -q_K & d^k_K & \cdots & d^l_K \\ -q^t_{K'} & -q^t_K & c_{00} & & c^t_K & \\ (d^k_{K'})^t & (d^k_K)^t & & & & \\ \vdots & \vdots & c_K & & C_K & \\ (d^l_{K'})^t & (d^l_K)^t & & & & \end{bmatrix}.$$

Since the product $EE^{-1} = I$, the identity matrix, the following conditions must be satisfied:

(1) $d^h_{K'} = 0$, $h = k,\ldots, l$; $d^h_K = e^{h-k+1}$, the $(h-k+1)$st unit vector in $\mathscr{R}^{|K|}$, $h = k,\ldots, l$.

(2) $C_K = 0$.

Since the matrix $S_{K'K'}$ is singular, the cofactor of b_{00} is zero and hence $c_{00} = 0$. Similarly, all the elements of b_K have cofactors equal to zero, since the minor obtained by deleting the row and the column intersecting at an element of b_k contains a zero row (or a zero column).

Since $EE^{-1} = I$, $c_{00} = 0$ and $c^t_K = 0$, $q^t_{K'}S_{K'K'} = 0$ and $q^t_{K'}\bar{v}_{K'} = 1$. Since $[\bar{S}_{K'K'}]^L = [\bar{p}_{K'}]$, $q_{K'} = \alpha\bar{p}_{K'}$, for some $\alpha \in \mathscr{R}$. Since $q_{K'}\bar{v} = 1 = \bar{p}_{K'}\bar{v}_{K'}$,

$q_{K'} = \bar{p}_{K'}$. Furthermore, since $-q^t_{K'}\bar{R}_{K'} - q^t_K = 0$, $q^t_K = -\bar{p}^t_{K'}\bar{R}_{K'} = \bar{p}^t_K$ by assumption. The matrix E^{-1} can now be written as follows:

$$E^{-1} = \begin{bmatrix} \bar{U}''_{K'K'} & \bar{U}''_{K'K} & -p_{K'} & 0 \\ \bar{U}''_{KK'} & \bar{U}''_{KK} & -p_K & I \\ -\bar{p}^t_{K'} & -\bar{p}^t_K & 0 & 0 \\ 0 & I & 0 & 0 \end{bmatrix}.$$

It can now be shown that, given any $\pi_K \in \mathbb{R}^K_+$, by appropriately choosing B_K we can guarantee that the matrix

$$\bar{U}'' = \begin{bmatrix} \bar{U}''_{K'K'} & U''_{K'K} \\ U''_{KK'} & \bar{U}''_{KK} \end{bmatrix}$$

is negative definite on the orthogonal complement of the vector $(\bar{p}^t_{K'}, \pi^t_K)$, $[(\bar{p}^t_{K'}, \pi^t_K)]^{\perp}$. Consider the matrices

$$M = \begin{bmatrix} \bar{U}''_{K'K'} & \bar{U}''_{K'K} & 0 \\ \bar{U}''_{KK'} & \bar{U}''_{KK} & I \\ 0 & I & 0 \end{bmatrix}, \qquad N = \begin{bmatrix} \alpha\bar{p}_{K'}\bar{p}^t_{K'} & \alpha\bar{p}_{K'}\pi^t_K & 0 \\ \alpha\pi_K\bar{p}^t_{K'} & \alpha\pi_K\pi^t_K & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$P = \begin{bmatrix} \bar{S}_{K'K'} - \bar{v}_{K'}\bar{v}^t_{K'} & 0 & \bar{R}_{K'} - \bar{v}_{K'}b^t_K \\ 0 & 0 & I \\ \bar{R}^t_{K'} - b_K\bar{v}^t_{K'} & I & B_K - b_Kb^t_K \end{bmatrix} \equiv \begin{bmatrix} P_1 & 0 & P_2 \\ 0 & 0 & I \\ P^t_2 & I & P_3 \end{bmatrix},$$

$$Q = \begin{bmatrix} 0 & 0 & \alpha\bar{p}_{K'}(\bar{p}_K - \pi_K)^t \\ \alpha(\bar{p}_K - \pi_K)v^t_{K'} & 0 & \alpha(\bar{p}_K - \pi_K)b^t_K + \alpha\pi_K(\bar{p}_K - \pi_K)^t \\ 0 & 0 & 0 \end{bmatrix},$$

and observe that for $\alpha = (b_{00} - 1)$, $(M + N)P = I + Q$. Furthermore, since $\bar{v}_{K'} \notin [\bar{S}_{K'K'}]$, the matrix P_1 is negative definite. Let Y be the set of vectors defined by $Y = \{(y_1, y_2, y_3) \in \mathscr{R}^{l+1+|K|} \mid y_2 = z_3 - P_3y_3 - P^t_2y_1,\ z_3 \in [y_3]^{\perp}, (y_1, y_3) \neq 0\}$. For $y = (y_1, y_2, y_3) \in Y$, $\alpha = (b_{00} - 1)$, $b_K = 0$, $B_K = \delta I$, $y^tP^t(M + N)Py = y^t_1P_1y_1 - \delta y^t_3Iy_3 - 2\alpha y^t_3(\pi_K - \bar{p}_K)\bar{v}^t_{K'}y_1 + \alpha y^t_3(\pi_K - \bar{p}_K)(\pi_K - \bar{p}_K)^ty_3$. Consequently, there exists $\bar{\delta} > 0$ such that, for $\delta > \bar{\delta}$, $B_K = \delta I$, $b_K = 0$, $\alpha = (b_{00} - 1)$, $y^tP^t(M + N)Py < 0$ for all $y \in Y$. *Consider the set $\bar{Y}$ of vectors defined by*

$$\bar{Y} = \{(\bar{y}_1, \bar{y}_2, \bar{y}_3) \in \mathscr{R}^{l+1+|K|} \mid \bar{y}_2 \in [\bar{y}_3]^{\perp}, (\bar{y}_1, \bar{y}_2) \neq 0\},$$

and observe that, given $\bar{y} \in \bar{Y}$ there exists a unique $y \in Y$ such that $\bar{y} = Py$: Explicitly,

$$(y_1, y_2, y_3) = (P_1^{-1}(\bar{y}_2 - P_2 \bar{y}_2), \bar{y}_3 - P_3 y_3 - P_2^t y_1, \bar{y}_2).$$

But then, $\bar{y}^t(M+N)\bar{y} < 0$ for all $\bar{y} \in \bar{Y}$. Equivalently, the matrix M is negative definite on the set $Y^* = \bar{Y} \cap [\bar{p}_{K'}, \pi_K, 0]^{\text{L}}$. But this is equivalent to the negative definiteness of $\bar{U}''$ on $[p_{K'}, \pi_K]^{\text{L}}$: for given $(y_1, y_2) \in [\bar{p}_{K'}, \pi_K]^{\text{L}}$ we can find y_3 such that $y = (y_1, y_2, y_3) \in Y^*$, and, furthermore, $y^t My = (y_1^t, y_2^t)\, \bar{U}'' \binom{y_1}{y_2}$. To complete the argument to the effect that Proposition (2) gives a set of not only necessary but also infinitesimally sufficient conditions, it suffices to construct an agent (u, w), satisfying Assumptions (1) and (2), and such that $D^2 u(\bar{x}^K) = \bar{U}''$, $Du(\bar{x}^K)^t = (p_{K'}^t, \pi_K^t)$, and $(\bar{x}^K + w) \in X$. But this is evident

Remark. To satisfy the system of equations (1) I have set $\lambda = 1$. Furthermore, since the vector π_K was chosen arbitrarily, no contradiction can arise concerning the sign of the multipliers μ_h, $h \in K$, when the equality constraints in (1) are replaced by the corresponding inequalities.

Remark. In the proof, b_{00} was left unspecified. By setting $b_{00} = 1$ (and hence $\alpha = 0$) we see that $M + N = M$. But then, for $b_{00} = 1$, $b_K = 0$, $B_K = \delta I$ the matrix M is negative definite on $\bar{Y}$. Equivalently the matrix $\bar{U}''$ is negative definite.

We have thus demonstrated the following:

Proposition 3. *Let $\bar{x}^K$ be a vector in $\mathscr{R}^{l+1}$, $\bar{s}$ a vector in $\mathscr{R}^{|K|}$, $\bar{p}$ a vector in $\mathscr{R}_+^{l+1}$, $\bar{S}_{K'K'}$ a $(|K'| \times |K'|)$ matrix, $\bar{v}_K$, a vector in $\mathscr{R}^{|K'|}$ and $\bar{R}_K$ a $(|K'| \times |K|)$ matrix, such that $\bar{x}_K^K = \bar{s}$, $\bar{p}^t x^K = 0$ and the following conditions are satisfied:*

(1) *$\bar{S}_{K'K'}$ is symmetric, negative semi-definite of rank $(|K'| - 1)$ and $\bar{p}_{K'}^t \bar{S}_{K'K'} = \bar{S}_{K'K'} \bar{p}_{K'} = 0$;*

(2) *$\bar{p}_{K'}^t \bar{v}_{K'} = 1$;*

(3) *$\bar{p}_{K'}^t \bar{R}_{K'} = -\bar{p}_K^t$.*

Then there exists an agent (u, w) satisfying Assumptions 1 and 2 whose constrained excess demand function $x^K(p, s)$ satisfies the following:

(a) *$x^K(\bar{p}, \bar{s}) = \bar{x}^K$;*

(b) *$D_{p_{K'}} x_{K'}^K(\bar{p}, \bar{s}) = \bar{S}_{K'K'} - \bar{v}_{K'}(\bar{x}_{K'}^K)^t$;*

(c) *$D_s x_{K'}^K(\bar{p}, \bar{s}) = \bar{R}_{K'}$.*

The following corollary follows immediately:

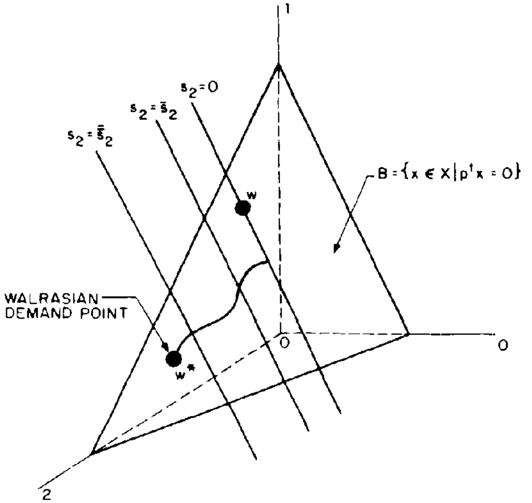


FIG. 1. The path (on B) to w^* is locally arbitrary.

COROLLARY 1. *The only restrictions imposed by Assumptions* 1 *and* 2 *on the spill-over matrix* $R_{K'}$ *is the adding up restriction*—$p^t_K R_{K'} = -p^t_K$.

It was argued in the introduction that the indeterminacy of the spill-over matrix $R_{K'}$ results from the lack of restrictions on the off-diagonal elements of the Hessian matrix $D^2 u$. I now want to pursue this line of reasoning for the purpose of determining the type of properties that, if satisfied by the Hessian matrix, have observable (qualitative) implications for the spill-over matrix. For simplicity I shall consider the case of three commodities, $h = 0, 1, 2$, and in particular the state of the market in which the consumer is constrained in the market for good 2, $K = \{2\}$ and unconstrained in the markets for goods 0 and 1, $K' = \{0, 1\}$.

By the definition of the matrices A and A^{-1},

$$R^t_{K'} D^2_{K'K'} u + D^2_{KK'} u + b_K p^t_{K'} = 0, \tag{4}$$

which can be written in terms of partial derivatives as follows: for all $h' \in K'$ and $h \in K$,

$$\sum_{h'' \in K'} \frac{\partial x_{h''}}{\partial s_h} \frac{\partial^2 u}{\partial x_{h''} \partial x_{h'}} + \frac{\partial^2 u}{\partial s_h \partial x'_h} + b_h p_{h'} = 0. \tag{5}$$

In the special case I am here considering, we get

$$\begin{aligned} \frac{\partial^2 u}{\partial x_0^2}\frac{\partial x_0}{\partial s_2} + \frac{\partial^2 u}{\partial x_0 \partial x_1}\frac{\partial x_1}{\partial s_2} + \frac{\partial^2 u}{\partial x_0\, \partial s} &= \frac{\partial \lambda^*}{\partial s_2} p_0, \\ \frac{\partial^2 u}{\partial x_0 \partial x_1}\frac{\partial x_0}{\partial s_2} + \frac{\partial^2 u}{\partial x_1^2}\frac{\partial x_1}{\partial s_2} + \frac{\partial^2 u}{\partial x_1\, \partial s_2} &= \frac{\partial \lambda^*}{\partial s_2} p_1. \end{aligned} \tag{5'}$$

Multiplying both sides of the first equation by $\partial u/\partial x_1$ and of the second equation by $\partial u/\partial x_0$, and subtracting the second equation from the first we get that

$$\frac{\partial x_0}{\partial s_2}\frac{\partial \sigma_{0,1}}{\partial x_0} + \frac{\partial x_1}{\partial s_2}\frac{\partial \sigma_{0,1}}{\partial x_1} + \frac{\partial \sigma_{0,1}}{\partial s_2} = 0, \tag{6}$$

where $\sigma_{0,1}$ is the marginal rate of substitution between goods 0 and 1, or $(\partial u/\partial x_0)/(\partial u/\partial x_1)$. Equation (6) is precisely the equation we want. The spill over matrix is now the vector $(\partial x_0/\partial s_2, \partial x_1/\partial s_2)$. Qualitative restrictions on the spill over effects can be derived by restricting the term $\partial\sigma_{0,1}/\partial s_2$, which captures the effects of a change in the consumption of good 2 on the marginal rate of substitution between goods 0 and 1.

3. EXTENSIONS: CONSTRAINED PRODUCTION DECISIONS

The results which I derived in the previous section can be immediately extended to the theory of production decisions. Here I shall only sketch the argument.[11]

The production plan of a producer is denoted by $y \in \mathbb{R}^{l+1}$—a commodity, h, is an output (resp. an input) if $y_h > 0$ (resp. if $y_h < 0$). The production possibilities of the firm are described by a function f: a plan y is feasible if and only if $f(y) \leqslant 0$. Given a price system p the firm chooses y so as to maximize $\pi = p^t y$ subject to the quantity constraints $y_h = s_h$, for $h \in K$. Since the problem degenerates if the firm is constrained in all markets, I shall assume that $K = \{k,..., l\}$ for some $k \geqslant 1$. The interpretation here of the 0th commodity is, of course, different from its interpretation in the theory of consumption. I shall assume that the function f is twice continuously differentiable, and strictly increasing, $f(0) = 0$, the production set $\{y \mid f(y) \leqslant 0\}$ is convex, and the hypersurface $\{y \mid 0 = f(y),\ y_h = s_h,\ h = k,..., l\}$ has nowhere vanishing Gaussian curvature. The first order necessary and sufficient conditions for a constrained maximum then becomes

---

[11] Here I follow, more or less, the analogous sketch in the Appendix in Laroque (1976).

$$\begin{aligned} p_{K'} - \lambda^* D_{K'} f(y^*) &= 0, \\ p_K - \lambda^* D_K f(y^*) - \mu^* &= 0, \\ f(y^*) &= 0, \\ y^*_K &= s, \end{aligned} \tag{6}$$

for some λ^*, μ^*_h, $h \in K$. To determine the Jacobian of the function $y^K(p, s)$ we totally differentiate (4):

$$\begin{bmatrix} \lambda^* D^2_{K'K'} f & \lambda^* D^2_{K'K} f & D_{K'} f & 0 \\ \lambda^* D^2_{KK'} f & \lambda^* D^2_{KK} f & D_K f & I \\ (D_{K'} f)^t & (D_K f)^t & 0 & 0 \\ 0 & I & 0 & 0 \end{bmatrix} \begin{bmatrix} dy_{K'} \\ dy_K \\ d\lambda^* \\ d\mu^* \end{bmatrix} = \begin{bmatrix} dp_{K'} \\ dp_K \\ 0 \\ ds \end{bmatrix}.$$

Setting

$$A = \begin{bmatrix} \lambda^* D^2_{K'K'} f & \lambda^* D^2_{K'K'} f & D_{K'} f & 0 \\ \lambda^* D^2_{KK'} f & \lambda^* D^2_{KK} f & D_K f & I \\ (D_{K'} f)^t & (D_K f)^t & 0 & 0 \\ 0 & I & 0 & 0 \end{bmatrix},$$

we see that the function $y^K(p, s)$ is completely described locally by the matrix

$$A^{-1} = \begin{bmatrix} S_{K'K'} & 0 & v^{K'} & R_{K'} \\ 0 & 0 & 0 & I \\ v^t_{K'} & 0 & b & b^t_K \\ R^t_{K'} & I & b_K & B_K \end{bmatrix}.$$

The matrix $R_{K'}$ is—as in the problem of the consumer—the spill-over matrix. The analogue of Propositions 1, 2, and 3 and of Corollary 1 can now be easily derived.

REFERENCES

BENASSY, J. P. (1975), Neokeynesian disequilibrium theory in a monetary economy, *Rev. Econ. Studies* **42** (4), No. 132, October, 503–524.

CLOWER, R. W. (1965), The Keynesian counterrevolution: A theoretical appraisal, *in* "The Theory on Interest Rates," (F. H. Hahn and F. R. R. Brechling, Eds.), Macmillan Co., New York.

DEBREU, G. (1972), Smooth preferences, *Econometrica* **40**, No. 4, July, 603–616.

DIAMOND, P. A. AND YAARI, M. (1972), Implications of the theory of rationing for consumer choice under uncertainty, *Amer. Econ. Rev.* **62**, No. 3, June, 333–343.

DRÈZE, J.(1975), Existence of an exchange equilibrium under price rigidities, *Internat. Econ. Rev.* **16**, 301–320.

GEANAKOPLOS, J. AND POLEMARCHAKIS, H. M. (1980), On the disaggregation of excess demand functions, *Econometrica* **48**, 315–332.

ITO, T. (1979), Methods of estimation for multi-market disequilibrium models, *Econometrica* **48**, No. 1, 97–126.

LAROQUE, G. (1979), The fixed price equilibria: Some results in local comparative statics, *Econometrica* in press.

LAROQUE, G. (1978), On the local uniqueness of the fixed price equilibria, mimeo, Institut National de la Statistique et des Etudes Economiques, April.

MALINVAUD, E. (1977), "The Theory of Unemployment Reconsidered," Halsted, New York.

MANTEL, R. (1977), "Implications of Microeconomic Theory for Community Excess Demand Functions," Cowles Foundation Discussion Paper No. 451.

POLEMARCHAKIS, H. M. (1979), Is there an income effect, *J. Econ. Theory* **21**, 380–388.

POLEMARCHAKIS, H. M., Expectations, demand and observability, *Econometrica*, in press.

TOBIN, J. (1952), A survey of the theory of rationing, *Econometrica* **20**, October, 521–553.

WIESMETH, H. (1977), Regular competitive equilibria in disequilibrium economies, mimeo, University of Hamburg, August.

JOURNAL OF ECONOMIC THEORY 25, 338-352 (1981)

Strategic Voting under Minimally Binary Group Decision Functions*

I. MACINTYRE AND PRASANTA K. PATTANAIK

Department of Economics, University of Birmingham, Birmingham B15 2TT, United Kingdom

Received July 4, 1979; revised August 14, 1980

The results on manipulability of not necessarily single-valued group decision functions are considerably more tentative than similar results for single-valued group decision functions.[1] The main reason is that when the set of outcomes may contain more than one alternative, non-obvious behavioural assumptions are needed to determine the individuals' preferences over possible sets of outcomes. The assumption of maximin behaviour and the assumption of expected utility maximization have been often used in this context.[2] However, these are rather restrictive assumptions which limit the scope of the theorems proved. In contrast, Barbera [3] proves an interesting result with a weak behavioural assumption. Barbera assumes that if an individual prefers x to y then he must prefer the outcome set $\{x\}$ to the outcome set $\{x, y\}$, and he must prefer the outcome set $\{x, y\}$ to the outcome set $\{y\}$. Given this very weak assumption he shows that every binary group decision function which is non-oligarchical and which satisfies Arrow's [1] condition of Citizen's Sovereignty must be manipulable.

While this result of Barbera is elegant with respect to the behavioural assumption used, it is extremely limited in scope in that it applies only to binary group decision functions (i.e., group decision functions under which for any given set of individual orderings, the choices from different sets of alternatives can be rationalized by a weak preference relation for the group). As is well known, the class of binary group decision functions is quite small, especially if the number of alternatives is sufficiently large.[3] Our paper seeks

---

* We are grateful to P. Dasgupta, R. Deb, P. Hammond, T. Gorman, and two anonymous referees for valuable comments on earlier versions of the paper.

[1] The well-known papers of Gibbard [7] and Satterthwaite [11] discuss the manipulability of single-valued group decision functions. Barbera [3, 4], Gardenfors [6], Gibbard [8], Kelly [9], Pattanaik [10], and Sengupta [15] are some of the writers who discuss manipulability of group decision functions which are not necessarily single-valued.

[2] For example, Sengupta [15] and in many cases Pattanaik [10] use the maximin assumption while Gibbard [8] uses the assumption of expected utility maximization.

[3] See Blau and Deb [2].

0022-0531/81/060338-15$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

to relax this assumption of binariness so as to retain the notion of pairwise comparisons only in a very minimal sense, while continuing to use behavioural assumptions which, though somewhat stronger than Barbera's assumption, are still intuitively extremely weak. In particular, we assume that for every individual the outcome set A is better than the outcome set B if for him every alternative in A is at least as good as every alternative in B, and some alternative in A is strictly better than some alternative in B.[4] Given this assumption it is shown that every group decision function which is based on pairwise comparisons in a very minimal sense and which does not allow any single individual to have any veto, must be manipulable, provided all possible weak orderings over the alternatives are permissible for individuals. On the other hand it is shown that if only strict orderings are permissible for individuals, then it is possible to construct a fairly wide class of "democratic" decision rules based on pairwise comparisons, which will be strictly non-manipulable under assumptions consistent with the behavioural rule referred to earlier.

In Section 1 we introduce our notation and some definitions. The main results are proved in Section 2. In Section 3 we comment on the significance of some of the assumptions.

1. THE NOTATION AND SOME DEFINITIONS

Let X be the set of alternatives and $N = \{1, 2, \ldots, n\}$ be the finite set of *individuals*. We assume that $|N| \geqslant 2$. In general we assume that $|X| \geqslant 3$; however, in some of the theorems we shall assume that $|X| \geqslant 4$. A non-empty subset of X is called an *issue* and a non-empty subset of N is called a *coalition*.

Let J be the set of all possible orderings over X and let J^0 be the set of all possible linear (or strict) orderings over X. For every $i \in N$, we have a given subset S_i of J, to be called the set of permissible orderings of individual i. The interpretation of S_i is as follows. S_i is the set of orderings over X which can figure as the true or sincere preference ordering of i; also S_i is the set of orderings which i can possibly express. Unless otherwise specified, we assume that $S_i = J$ for all $i \in N$ (in Section 3 we consider the case where $S_i = J^0$ for every $i \in N$). $S_1 \times S_2 \times \cdots \times S_n$ will be indicated by S. The elements $(R_1, \ldots, R_n)$, $(R'_1, \ldots, R'_n)$, etc., of S will be called *situations* and will be indicated by s, s', etc. $R_1, \ldots, R_n$ have the usual interpretation as orderings of the individuals. Similarly for $R'_1, \ldots, R'_n$. The true or sincere preference ordering of an individual i ($i \in N$) will be indicated by $\bar{R}_i$. $\bar{s} = (\bar{R}_1, \ldots, \bar{R}_n)$ will be called the *sincere situation*. Corresponding to R_i, P_i and I_i are the strict preference and indifference relations defined in the usual fashion. Similarly $\bar{P}_i$ and $\bar{I}_i$ correspond to $\bar{R}_i$ and so on. For all coalitions L and all $x, y \in X$, $xR_L y$ iff ($xR_i y$ for all $i \in L$). Similarly we have $xP_L y$, $xI_L y$, etc.

---

[4] Given the assumption that individual preference are orderings, this can be shown to be equivalent to the following rule: A is better than B if every alternative in A is at least as good as every alternative in B and either some alternative in A is better than every alternative in B or some alternative in B is worse than every alternative in A. See Pattanaik [10].

DEFINITION 1. A *group decision function* (GDF) is a function f which for every situation s and every issue A, specifies exactly one non-empty subset E of A. We write $E = f(s, A)$ and call E the set of outcomes yielded by f given situation s and issue A.

For all $x, y \in X$ and for all $s \in S$, xRy iff ($x \in f(s, \{x, y\})$); xPy iff (xRy and $\sim yRx$); and xIy iff (xRy and yRx). Similarly we define R', P' and I' corresponding to s' and so on. For any given issue A, $C(A, R) = \{x \in A \mid xRy$ for all $y \in A\}$.

We assume that for every $i \in N$ and every possible sincere ordering $\bar{R}_i$ of i, $\bar{R}_i$ generates a binary weak preference relation $[\bar{R}_i \geqslant]$ over the set of all possible non-empty subsets of X. The binary relation $[\bar{R}_i \geqslant]$ is to be interpreted as the ith individual's preferences over alternative conceivable sets of outcomes. $[\bar{R}_i >]$ is the strict preference relation corresponding to $[\bar{R}_i \geqslant]$.

Consider the following three assumptions due to Barbera [3], Kelly [9], and Gardenfors [6], respectively.

ASSUMPTION 1 (Barbera [3]). *For all $i \in N$, and for all non-empty subsets G, G', and G'' of X, $G[\bar{R}_i >]G'$ and $G'[\bar{R}_i >]G''$ if for some $x, y \in X$, $G = \{x\}$; $G' = \{x, y\}$; $G'' = \{y\}$; and $x\bar{P}_i y$.*

ASSUMPTION 2 (Kelly [9]). *For all $i \in N$ and for all non-empty subsets G and G' of X, $G[\bar{R}_i >]G'$ if (for all $x \in G$ and for all $x' \in G'$, $x\bar{R}_i x'$) and (for some $x \in G$ and for some $x' \in G'$, $x\bar{P}_i x'$).*

ASSUMPTION 3 (Gardenfors [6]). *For all $i \in N$, and for all non-empty subsets G and G' of X, $G[\bar{R}_i >]G'$ if any of the following conditions is fulfilled:*

(i) *$G \subset G'$; for all $x \in G$ and all $y \in (G' - G)$, $x\bar{R}_i y$; and for some $x \in G$ and some $y \in (G' - G)$, $x\bar{P}_i y$;*

(ii) *$G' \subset G$; for all $x \in (G - G')$ and all $y \in G'$, $x\bar{R}_i y$; and for some $x \in (G - G')$ and some $y \in G'$, $x\bar{P}_i y$.*

(iii) *Neither $G \subset G'$ nor $G' \subset G$ nor $G = G'$; for all $x \in (G - G')$ and all $y \in (G' - G)$, $x\bar{R}_i y$; and for some $x \in (G - G')$ and some $y \in (G' - G)$, $x\bar{P}_i y$.*

Assumption 1 is the weakest of the three. Assumption 2, though somewhat stronger than Assumption 1, is intuitively extremely plausible. All that Assumption 2 says is that if everything in G is at least as good as everything in G' and something in G is strictly better than something in G', then G is better than G'. We shall mostly use Assumption 2 in this paper. Assumption 3 is much stronger than Assumption 2. The intuitive appeal of Assumption 3 is not quite clear. For example, if $x_1 \bar{P}_i x_2 \bar{P}_i \cdots \bar{P}_i x_{10}$, then under Assumption 3, $\{x_1, x_2, \ldots, x_9\} [\bar{R}_i >] \{x_1, x_{10}\}$ while intuitively this is not a compulsive conclusion. (In fact this may be inconsistent with expected utility maximization depending on the utility function representing $\bar{R}_i$ and the probabilities attached to the alternatives in the two sets.)

Replacing "if" by "only if" in Assumptions 1, 2 and 3 we get Assumptions 1′, 2′ and 3′, which are in descending order of logical strength. Assumptions 1′, 2′ and 3′ lack plausibility and are extremely strong assumptions to make.

We now introduce the familiar notion of manipulability. For all $s, s' \in S$ and for all coalitions L, s and s' are said to be *L-variant* iff $[(R_i \neq R_i'$ for all $i \in L)$ and (for all $j \in (N - L)$, $R_j = R_j')]$. (When L consists of a single individual, say k, we write "k-variant" instead of writing "$\{k\}$-variant.")

DEFINITION 2. Let f be a GDF. Given an issue A and the sincere situation $\bar{s}$, $s \in S$ is said to be a *strict equilibrium* (resp. an *equilibrium*) iff there does not exist (s', L) such that $s' \in S$; L is a coalition (resp. L is a coalition with exactly one member); s and s' are L-variant; and for all $i \in L$, $f(s', A) [\bar{R}_i >] f(s, A)$. (If such (s', L) exists, we say that (s', L) is a *threat* to s given the issue A and sincere situation $\bar{s}$; when L contains only one individual, say k, we write (s', k) instead of writing $(s', \{k\})$ to indicate the threat.) f is *strictly non-manipulable* (resp. *non-manipulable*) iff for every possible issue, every possible sincere situation $\bar{s}$ ($\bar{s} \in S$) is a strict equilibrium (resp. equilibrium).

Our next definition introduces several properties of a GDF.

DEFINITION 3. Let s and s' be any two situations in S, and let x and y be any two alternatives. Let A be any issue and let f be the GDF.

(3.1) *Binariness* (B): $C(A, R) = f(s, A)$.

(3.2) *Minimal Binariness* (MB): If $C(A, R) \neq \varnothing$, then

$$C(A, R) \subseteq f(s, A) \subseteq \{x \in A \mid \sim\exists y \in C(A, R) : yPx\}.$$

(3.3) *Schwartz's Rule* (SZ): $f(s, A) = \bigcup_{B \in \mathscr{B}} B$, where $B \in \mathscr{B}$ iff (i) $B \subseteq A$; and (ii) for all $a, b \in X$ if $a \in B$ and $b \in (A - B)$, then $\sim bPa$; and (iii) there does not exist any proper subset of B which satisfies (i) and (ii).

(3.4) *Limited Independence of Irrelevant Alternatives* (LIND): If, for all $i \in N$, $[(xR_i y \text{ iff } xR'_i y)$ and $(yR_i x \text{ iff } yR'_i x)]$ then $[(xRy \text{ iff } xR'y)$ and $(yRx \text{ iff } yR'x)]$.

(3.5) LIND*: Suppose for all $i \in N$, if $xI_i y$, then $R_i = R'_i$ and if $\sim xI_i y$, then $[(xR_i y \text{ iff } xR'_i y)$ and $(yR_i x \text{ iff } yR'_i x)]$. Then $[(xRy \text{ iff } xR'y)$ and $(yRx \text{ iff } yR'x)]$.

(3.6) Limited Monotonicity (LM): If $[$(for all $i \in N$ and all $a \in X$, $xP_i a \to xP'_i a$ and $xI_i a \to xR'_i a$) and (for all $i \in N$ and all $a, b \in (X - \{x\})$, $aR_i b$ iff $aR'_i b$)$]$, then $[xPy \to xP'y$ and $xIy \to xR'y]$.

(3.7) *Limited Absence of Individual Vetoes* (LAV): If $|\{i \in N | xP_i y\}| \geqslant |N| - 1$, then xPy.

(3.8) *Limited Sovereignty of Coalitions of at least k ($|N| \geqslant k \geqslant 1$) Individuals* ($\text{LSOV}_k$): Let L be any coalition with at least k members. Then for all $i \in L$, there exist $\hat{R}_i \in S_i$ such that for all $s \in S$, if $R_i = \hat{R}_i$ for all $i \in L$, then $x\hat{P}y$.

(3.9) *Limited Resoluteness* (LRS): (xPy or yPx).

B and MB are concerned with two alternative senses in which one can speak of a GDF being based on pairwise comparison of alternatives. B requires $f(s, A)$ to coincide with the set of R-greatest element in A. Since by definition, $f(s, A)$ must be non-empty, B implies that an R-greatest element in A must exist. As is well known this is an extremely restrictive requirement which excludes most democratic decision procedures especially when $|A|$ is sufficiently large.

MB is much weaker than B. First, unlike B, MB does not impose any restriction on $f(s, A)$ when there does not exist any R-greatest element in A. Secondly, when R-greatest elements in A do exist, MB does not require $f(s, A)$ to coincide necessarily with the set of R-greatest elements in A. If $C(A, R) \neq \varnothing$, MB only requires that $C(A, R)$ should be a subset of $f(s, A)$ which in turn should be a subset of the set of alternatives x such that x is not defeated by any element of $C(A, R)$ in a pairwise comparison under f. Intuitively speaking, all that MB requires is this: If at all there exists an all round winner (i.e., alternative which is not defeated in a pairwise comparison with any alternative in the issue) in the issue, then every such all round winner must figure in the set of outcomes and no alternative which is defeated in a pairwise comparison with some all round winner in the issue can possibly belong to the set of outcomes.

The importance of the fact that MB permits $C(A, R)$ to be empty is obvious. What may not be so obvious is the significance of the requirement that where $C(A, R)$ is non-empty, $C(A, R)$ should be a subset of, but not necessarily coincident with, $f(s, A)$. We clarify this below.

For any given issue A, let Q_A be the set of all one-to-one functions from $\{1, 2, \dots, |A|\}$ to A. Let f be any given GDF. For all $q \in Q_A$ and for all $s \in S$, let $A^1(q, s, f) = \{q(1)\}$ and for all integers i $(i = 1, 2, \dots, |A| - 1)$ let $A^{i+1}(q, s, f) = \bigcup_{x \in A^i(q,s,f)} f(s, \{q(i+1), x\}) - \{y \in A \mid y = q(i+1)$ and xPy for some $x \in A^i(q, s, f)\}$. Consider the class F of GDFs f such that for every issue A there exists $q \in Q_A$ such that for every situation s, $f(s, A) = A^{|A|}(q, s, f)$. Intuitively F is the class of GDFs under which choice from any given issue is based on a sequence of pairwise comparisons. ($q \in Q_A$ can be interpreted as specifying a sequence or "path" of pairwise comparison of alternatives in A.) It can be easily checked that for every issue A, every $s \in S$, and every $q \in Q_A$, $C(A, R) \subseteq A^{|A|}(q, s, f) \subseteq \{x \in A \mid {\sim}\exists y \in C(A, R) : yPx\}$. Therefore every $f \in F$ satisfies MB. However, many GDFs belonging to F violate the property that if $C(A, R) \neq \varnothing$, then $f(s, A)$ should be identical with $C(A, R)$. This is shown by Example 1. Example 2 illustrates certain other types of GDFs (not based on a sequence of pairwise comparisons) which, though satisfying MB, allow $C(A, R)$ and $f(s, A)$ to differ when $C(A, R) \neq \varnothing$.

Example 1. Suppose $X = \{x, y, z\}$. Let X itself be the issue. The agenda is to compare x and y first and then to compare the winners with z, and to declare the winners at this stage as the winners in the entire set of X. (In terms of our notation, let $q \in Q_X$ be such that $q(1) = x$; $q(2) = y$; and $q(3) = z$. Then $f(s, X) = X^3(q, s, f)$. Suppose $s \in S$ is such that we have (xPy and yPz and xIz). Then following the method mentioned we have $f(s, X) = \{x, z\}$ even though $C(X, R) = \{x\}$. It is clear that this is consistent with MB.

Example 2. Suppose $A = \{x_1, x_2, \dots, x_6\}$. Suppose for all $s \in S$ if $C(|C(\{x_1, x_2\}, R) \cup C(\{x_3, x_4\}, R) \cup C(\{x_5, x_6\}, R)|, R) \neq \varnothing$, then $f(s, A) = C(|C(\{x_1, x_2\}, R) \cup C(\{x_3, x_4\}, R) \cup C(\{x_5, x_6\}, R)|, R)$; and in all other cases $f(s, A) = A$. Suppose for all $y \in A$, $x_1 R y$; and we have ($x_1 P x_2$ and $x_3 P x_4$ and $x_5 P x_6$ and $x_1 P x_3$ and $x_5 P x_3$ and $x_1 I x_5$ and $x_2 P x_5$). Clearly $f(s, A) = \{x_1, x_5\}$ but $C(A, R) = \{x_1\}$. However, this is consistent with MB.

The well-known condition of Schwartz's Rule (SZ)[5] is stronger than the property MB.

Among the other properties introduced in Definition 3, LIND, LM, LAV, $\mathrm{LSOV}_{|N|-1}$ and LRS are familiar; LIND* is a slightly weaker version of LIND.

---

[5] See Schwartz [12, 13].

2. MANIPULABILITY OF MINIMALLY BINARY GROUP DECISION FUNCTIONS—THE CASE OF WEAK INDIVIDUAL ORDERINGS

In this section we shall assume that for every individual i, $S_i = J$. Given this assumption and given Assumption 2, we show that every minimally binary GDF satisfying $\text{LSOV}_{|N|-1}$ is manipulable if $|X| \geqslant 4$ and that every minimally binary GDF satisfying $\text{LSOV}_{|N|-1}$ and LIND is manipulable if $|X| \geqslant 3$. We first prove the following results.

LEMMA 1. *Let A be a given issue and let $s, s' \in S$ be such that $f(s, A) \neq f(s', A)$ where f is the* GDF. *Then there exist $i \in N$ and i-variant $\hat{s}, \tilde{s} \in S$ such that $[(\hat{R}_i = R_i)$ and $(\tilde{R}_i = R'_i)]$; [for all $j \neq i$ $(\hat{R}_j = R_j$ or $\hat{R}_j = R'_j)$ and $(\tilde{R}_j = R_j$ or $\tilde{R}_j = R'_j)]$; and $[f(\hat{s}, A) = f(s, A) \neq f(s, A)]$.*

Proof. Starting with s, replace the orderings R_i by R'_i $(i = 1,..., n)$, one at a time, thereby generating the finite sequence of situations $(s^1, s^2,..., s^n)$. [Thus $s^1 = (R'_1, R_2,..., R_n)$ and for all i $(i = 1,..., n-1)$, $s^{i+1} = (R^i_1, R^i_2,..., R^i_i, R'_{i+1}, R^i_{i+2},..., R^i_n)$]. Clearly $s^n = s'$. Let t be the smallest integer $(1 \leqslant t \leqslant n)$ such that $f(s^t, A) \neq f(s, A)$. Since $f(s^n, A) = f(s', A) \neq f(s, A)$, such t exists. If $t = 1$, then let $\hat{s} = s$ and $\tilde{s} = s^1$. If $t > 1$, then let $\hat{s} = s^{t-1}$ and $\tilde{s} = s^t$. Clearly $\hat{s}$ and $\tilde{s}$ constructed in this way satisfy all the conditions stipulated in the statement of the lemma. ∎

LEMMA 2. *Suppose a* GDF *f violates* LIND*. *Then f must be manipulable given Assumption* 2.

Proof. Since f violates LIND*, by definition of LIND* we have $x, y \in X$ and $\hat{s}, \tilde{s} \in S$ such that for all $i \in N$, $[(x\hat{P}_i y$ and $x\tilde{P}_i y)$ or $(y\hat{P}_i x$ and $y\tilde{P}_i x)]$ and $f(\hat{s}, \{x, y\}) \neq f(\tilde{s}, \{x, y\})$. Hence given Lemma 1, we have $x, y \in X$ and $\hat{s}, \tilde{s} \in S$ such that for some $i \in N$, $\hat{s}$ and $\tilde{s}$ are i-variant; $[(x\hat{P}_i y$ and $x\tilde{P}_i y)$ or $(y\hat{P}_i x$ and $y\tilde{P}_i x)]$; and $[f(\hat{s}, \{x, y\}) \neq f(\tilde{s}, \{x, y\})]$. Since $f(\hat{s}, \{x, y\}) \neq f(\tilde{s}, \{x, y\})$, either $(x\hat{P}y$ and $y\tilde{R}x)$ or $(y\hat{R}x$ and $x\tilde{P}y)$.

Suppose $(x\hat{P}y$ and $y\tilde{R}x)$. If $(x\hat{P}_i y$ and $x\tilde{P}_i y)$, then assuming $\tilde{s} = \bar{s}$, $(\hat{s}, i)$ is a threat to $\bar{s}$ and hence f is manipulable. If $(y\hat{P}_i x$ and $y\tilde{P}_i x)$, then assuming $\hat{s} = \bar{s}$, $(\tilde{s}, i)$ is a threat to $\bar{s}$ and hence f is manipulable.

The proof for the case where $(y\hat{R}x$ and $x\tilde{P}y)$ is similar to the one in the preceding paragraph. ∎

LEMMA 3. *If a* GDF *violates* LM, *then it must be manipulable given Assumption* 2.

For the proof of this result, the reader may refer to Pattanaik [10, pp. 62–63].

Let f be any given GDF. For any ordered pair of alternatives $x, y \in X$ and any coalition L, L is nearly decisive for x against y iff $x \neq y$ and for all $s \in S$, if ($xP_L y$ and $yP_{N-L}x$) then xPy. Let D_f be the set of all smallest nearly decisive coalitions (so that a coalition L belongs to D_f iff L is nearly decisive for some x against some y, and $|L| \leqslant |\hat{L}|$ for every coalition $\hat{L}$ which is nearly decisive for some $\hat{x}$ against some $\hat{y}$.) If f satisfies LAV, then D_f is clearly non-empty, and for all $L \in D_f$, $|L| \geqslant 2$ and $N - L \neq \varnothing$.

LEMMA 4. *Let Assumption 2 be satisfied. Let f be a minimally binary* GDF *satisfying* LAV *and let $L \in D_f$. Let L be nearly decisive for a against b ($a, b \in X$ and $a \neq b$). Then f is manipulable if at least one of the following conditions holds*:

(α) *there exist $s^0 \in S$, $c \in (X - \{a, b\})$, and $h \in L$ such that* [(*for all* $i \in (L - \{h\})$, $aP_i^0 bP_i^0 c$) *and* ($bP_{N-L}^0 cP_{N-L}^0 a$) *and* ($cP_h^0 aI_h^0 b$) *and* ($bR^0 a$)]; *or*

(β) *there exist $s^0 \in S$, $c \in (X - \{a, b\})$, and $g, h \in L$ such that* [($aI_h^0 bP_h^0 c$) *and* (*for all* $i \in (L - \{g, h\})$, $aP_i^0 bP_i^0 c$) *and* ($bP_{N-L}^0 cP_{N-L}^0 a$) *and* ($cP_g^0 aP_g^0 b$) *and* ($aP^0 b$)].

Proof. If f violates LIND* or LM, then f is manipulable by Lemmas 2 and 3. Hence in what follows we assume that f satisfies LIND* and LM.

I. Suppose (α) holds. Consider s^0 as specified in (α). By LAV we have bP^0c. We cannot have aP^0c. For, if aP^0c, then by LIND, $(L - \{h\}) \in D_f$, which contradicts the fact that $L \in D_f$. Ths cR^0a. Therefore (bR^0a and bP^0c and cR^0a). Let $A = \{a, b, c\}$. By MB, $f(s^0, A) = \{b\}$ or $f(s^0, A) = \{a, b\}$. Consider $s^1 \in S$ such that s^1 and s^0 are h-variant; for all $x, y \in A$, if $\{x, y\} \neq \{a, b\}$, then $xR_h^0 y$ iff $xR_h^1 y$; and $aP_h^1 b$. By LIND* comparing s^1 and s^0 we have (bP^1c and cR^1a); and given that L is nearly decisive for a against b, we have (aP^1b). If $c \in f(s^1, A)$, then assuming $s^0 = \bar{s}$, (s^1, h) will be a threat to $\bar{s}$ and hence f will be manipulable. If $f(s^1, A) \subseteq \{a, b\}$, then there are three possibilities: (i) $f(s^1, A) = \{a\}$; (ii) $f(s^1, A) = \{b\}$; or (iii) $f(s^1, A) = \{a, b\}$.

I(i). Suppose $f(s^1, A) = \{a\}$. Consider s^2 such that s^2 and s^1 are $(N - L)$-variant and $cP_{N-L}^2 bP_{N-L}^2 a$. By LIND*, comparing s^1 and s^2 we have (aP^2b and cR^2a). By the fact that $L \in D_f$, we have cR^2b. Hence by MB, $c \in f(s^2, A)$ and $f(s^2, A) \neq f(s^1, A)$. Hence by Lemma 1, there exist $\hat{s}, \bar{s} \in S$ and $k \in (N - L)$ such that $\hat{s}$ and $\bar{s}$ are k-variant; $f(\hat{s}, A) = \{a\} \neq f(\bar{s}, A)$; and ($b\hat{P}_k c\hat{P}_k a$). Assuming $\hat{s} = \bar{s}$ and given our Assumption 2, $(\bar{s}, k)$ is a threat to $\bar{s}$ and hence f is manipulable.

I(ii). If $f(s^1, A) = \{b\}$, then consider $s^3 \in S$ such that s^3 and s^1 are $(N - L)$-variant and $bP_{N-L}^3 aP_{N-L}^3 c$. By LIND* and by AV, we have (aP^3b and bP^3c and aP^3c). Hence $f(s^3, A) = \{a\}$ and $f(s^1, A) \neq f(s^3, A)$. Then by Lemma 1 we can construct $\hat{s}$ and $\bar{s}$ such that $\hat{s}$ and $\bar{s}$ are k-variant for some $k \in (N - L)$; $f(\hat{s}, A) = \{b\} \neq f(\bar{s}, A)$; and $b\tilde{P}_k a\tilde{P}_k c$. Assuming $\bar{s} = \bar{s}$, $(\hat{s}, k)$ is a threat to $\bar{s}$ and hence f is manipulable.

I(iii). Suppose $f(s^1, A) = \{a, b\}$. Then consider s^3 as constructed above. $f(s^1, A) \neq f(s^3, A) = \{a\}$. Hence by Lemma 1, we can construct $\hat{s}$ and $\tilde{s}$ such that $\hat{s}$ and $\tilde{s}$ are k-variant for some $k \in (N - L)$; for all $i \neq k$, $(\tilde{R}_i = R_i^1$ or $\tilde{R}_i = R_i^3)$; $(b\hat{P}_k c\hat{P}_k a$ and $b\tilde{P}_k a\tilde{P}_k c)$; and $f(\hat{s}, A) = \{a, b\} \neq f(\tilde{s}, A)$. By LIND*, comparing $\tilde{s}$ and s^3 we have $(a\tilde{P}b$ and $b\tilde{P}c)$. Since $f(\tilde{s}, A) \neq \{a, b\}$, $f(\tilde{s}, A)$ can assume any one of six possible values: $\{a\}$ or $\{c\}$ or $\{a, c\}$ or $\{b\}$ or $\{b, c\}$ or $\{a, b, c\}$. If $f(\tilde{s}, A)$ is equal to $\{a\}$ or $\{c\}$ or $\{a, c\}$ then assuming $\tilde{s} = \bar{s}$, $(\hat{s}, k)$ will be a treat to $\bar{s}$ and hence f will be manipulable. If $f(\tilde{s}, A) = \{b\}$, then assuming $\hat{s} = \bar{s}$, $(\tilde{s}, k)$ is a threat to $\bar{s}$ and hence f is manipulable. Suppose $f(\tilde{s}, A) = \{b, c\}$ or $f(\tilde{s}, A) = \{a, b, c\}$. Then given MB, $(a\tilde{P}b$ and $b\tilde{P}c$ and $a\tilde{R}c)$ cannot hold. Since $(a\tilde{P}b$ and $b\tilde{P}c)$, it follows that we must have $(a\tilde{P}b$ and $b\tilde{P}c$ and $c\tilde{P}a)$. Note that by construction, $\tilde{R}_h = R_h^1$. Consider $s^4 \in S$ such that s^4 and $\tilde{s}$ and h-variant and $R_h^4 = R_h^0$. Comparing s^4 with s^0 and $\tilde{s}$, we have $(bR^4a$ and bP^4c and $cP^4a)$ by LIND*. Hence $b \in f(s^4, A) \subseteq \{a, b\}$. Assume $s^4 = \bar{s}$; since by assumption $f(\tilde{s}, A)$ is equal to $\{b, c\}$ or $\{a, b, c\}$, it is then clear that $(\tilde{s}, h)$ is a threat to $\bar{s}$ and hence f is manipulable in these cases also.

II. Suppose (β) holds. Consider s^0 as specified in (β) and let $A = \{a, b, c\}$. As in Step I, we must have $(bP^0c$ and $cR^0a)$. Since aP^0b (by assumption), we have $(aP^0b$ and bP^0c and $cR^0a)$. If $c \in f(s^0, A)$, then consider s^1 such that s^1 and s^0 are h-variant and $bP_h^1aP_h^1c$. By LIND* and by the fact that $L \in D_f$, we have $(bR^1a$ and bP^1c and $cR^1a)$. Hence by MB, $f(s^1, A) \subseteq \{a, b\}$. Then assuming $s^0 = \bar{s}$, (s^1, h) will be a threat to $\bar{s}$ and hence f will be manipulable. If $c \notin f(s^0, A)$, then $f(s^0, A) \subseteq \{a, b\}$. There are three possibilities: $f(s^0, A) = \{a\}$, or $f(s^0, A) = \{b\}$, or $f(s^0, A) = \{a, b\}$.

If $f(s^0, A) = \{a\}$ or $f(s^0, A) = \{b\}$, then the proof is exactly similar to the proof in Steps I(i) and I(ii), respectively.

If $f(s^0, A) = \{a, b\}$, we proceed exactly in the same way as in Step I(iii) except for one difference. In Step I(iii), in the case where $f(\tilde{s}, A)$ was $\{b, c\}$ or $\{a, b, c\}$ we constructed $s^4 \in S$ such that assuming $s^4 = \bar{s}$, $(\tilde{s}, h)$ turned out to be a threat to $\bar{s} = s^4$. In the corresponding stage here we have to assume that $\tilde{s} = \bar{s}$ and then (s^4, h) will turn out to be a threat to $\bar{s} = \tilde{s}$. Except for this difference the proof here remains exactly similar to the proof in Step I(iii). ∎

We now prove our Theorems 1 and 2.

THEOREM 1. *Let Assumption* 2 *be satisfied and let* $|X| \geqslant 4$. *Let* f *be a minimally binary* GDF *satisfying* $\text{LSOV}_{|N|-1}$. *Then* f *is manipulable.*

Proof. If f violates LIND* or LM, f will be manipulable (by Lemmas 2 and 3). So we consider only the case where f satisfies LIND* and LM. Given LIND* and LM, it can be checked that $\text{LSOV}_{|N|-1}$ implies LAV. So in the rest of the proof we assume that f satisfies LAV and MB.

Let $L \in D_f$. Let L be nearly decisive for a against b ($a, b \in X$ and $a \neq b$). Let $c, d \in X$ be such that a, b, c, d are all distinct. By LAV, $|L| \geqslant 2$. Let g and h be two distinct individuals in L. Construct s such that

$$\begin{array}{ll} \text{for all } i \in (L - \{g, h\}), & aP_i bP_i cP_i d; \\ & dP_g aP_g bP_g c; \\ & cP_h aI_h bP_h d; \\ & \\ & bP_{N-L} dP_{N-L} cP_{N-L}\, a. \end{array}$$

Either bRa or aPb. If bRa, then considering the triple (a, b, c) we know that by Lemma 4, f is manipulable. On the other hand, if aPb, then considering the triple (a, b, d), we know that by Lemma 4 again, f is manipulable. Thus f is manipulable in all cases. ▮

If one is prepared to add LIND to the properties MB and $\text{LSOV}_{|N|-1}$ in the statement of Theorem 1, then the condition, $|X| \geqslant 4$, can be weakened to the condition that $|X| \geqslant 3$.

THEOREM 2. *Let Assumption 2 be satisfied and let* $|X| \geqslant 3$. *Let* f *satisfy* LIND, MB *and* $\text{LSOV}_{|N|-1}$. *Then* f *is manipulable.*

Proof. As in the proof of Theorem 1 we need consider only the case where f satisfies LM. Since given LM and LIND, $\text{LSOV}_{|N|-1}$ implies LAV, in the rest of the proof we assume that f satisfies LAV.

Let $L \in D_f$. Let L be nearly decisive for a against b ($a, b \in X$ and $a \neq b$). Let $c \in (X - \{a, b\})$. Let $g, h \in L$ ($g \neq h$). Let s be any situation such that $(aI_h b)$ and (for all $i \in (L - \{h\})$, $aP_i b$) and $(bP_{N-L}\, a)$. We must have either bRa or aPb.

If bRa, then construct $s^0 \in S$ such that $[$(for all $i \in (L - \{h\})$, $aP_i^0 bP_i^0 c$) and $(bP_{N-L}^0 cP_{N-L}^0 a)$ and $(cP_h^0 aI_h^0 b)]$. By LIND, we have $bR^0 a$. Hence by Lemma 4, f is manipulable.

If aPb, then construct $s^0 \in S$ such that $[(aI_h^0 bP_h^0 c)$ and (for all $i \in (L - \{g, h\})$, $aP_i^0 bP_i^0 c$) and $(bP_{N-L}^0 cP_{N-L}^0 a)$ and $(cP_g^0 aP_g^0 b)]$. By LIND, we have $aP^0 b$. Hence by Lemma 4, f is manipulable. ▮

Note that if at all Theorem 1 had assumed LIND*, then the cost involved in changing to LIND in Theorem 2 would be only marginal. However, Theorem 1 does not assume either LIND* or LIND; hence introduction of LIND in Theorem 2 constitutes a significant additional restriction. It is true that LIND* played a crucial role in the proof of Theorem 1 but it was not necessary to assume LIND* in Theorem 1 since LIND* is a necessary condition for non-manipulability of a GDF under Assumption 2. LIND also plays an equally crucial role in the proof of Theorem 2. However, it is not

TABLE I

Assumptions under Which Three Alternative Theorems Show That the GDF Is Manipulable

	(1)	(2)	(3)	(4)	(5)
Theorem 1 in this paper	$\lvert X\rvert \geqslant 4$ and $\lvert N\rvert \geqslant 2$	Assumption 2	MB	LSOV$_{\lvert N\rvert}$	—
Barbera's theorem	$\lvert X\rvert \geqslant 3$ and $\lvert N\rvert \geqslant 2$	Assumption 1	B	LSOV$_{\lvert N\rvert}$ and Absence of Oligarchy	—
Theorem of Gardenfors	$\lvert X\rvert \geqslant 3$ and $\lvert N\rvert \geqslant 3$	Assumption 3	—	Condorcet criterion	Anonymity Neutrality

possible to appeal to any "necessity" result to discard LIND in Theorem 2 since given our assumption that all *weak* orderings are permissible for an individual (i.e., $S_i = J$ for all $i \in N$), LIND is *not* a necessary condition for non-manipulability under Assumption 2 (see Example 5.1 in Pattanaik [10, pp. 78–79]).

Theorem 1 is compared with the central theorems of Barbera [3] and Gardenfors [6] in Table I, where we list the assumptions under which each of these theorems shows that f is manipulable.[6] While some assumptions of Barbera are weaker than corresponding assumptions of our Theorem 1 (see columns (1), (2) and (4) in Table I), this is achieved at a very high cost since his property B rules out most democratic decision rules based on pairwise comparison while MB admits most of them. Gardenfors does not use either B or MB but his other assumptions—Condorcet criterion, anonymity and neutrality—are rather strong.

It may also be useful to comment on the relationship between our Theorem 1 and Theorem 1 of Barbera [4] which investigates manipulability (under Assumption 1) of not necessarily single-valued GDFs in a somewhat different framework. The major difference lies in that Barbera does not use any of the "binariness condition" but postulates a property called "positive reponsiveness." This property requires that if for a given situation, x is one of many outcomes, and if x moves up vis-a-vis some alternatives in some individual's orderings (the orderings remaining intact otherwise), then in the resulting new situation x must be the only outcome. Note that while discarding all "binariness conditions" is a gain, positive responsiveness is a restrictive condition violated by most decision rules covered by the wide class of GDFs figuring in our theorem. As can be easily checked, positive responsiveness is violated by every GDF which satisfies MB and LAV in addition to the following property: For all $a, b \in X$ and for all situations s, if $|\{i \in N \mid aP_i b\}| = |\{i \in N \mid bP_i a\}|$, then aIb.

---

[6] Certain properties mentioned in Table I have not been defined earlier. We define them here. Let s and s' be any two situations; A be any issue; σ be any one-to-one function from N to N; and ψ be any one-to-one function from X to X. A GDF f satisfies: (1) *Absence of Oligarchy* iff there does not exist a unique coalition L such that for all distinct $a, b \in X$, if aP_Lb, then aPb, and if aP_ib for some $i \in L$, then aRb; (2) *Condorcet Criterion* iff for all $a \in A$, if $[|\{i \in L \mid aP_ib\}| > N/2$ for all $b \in (A - \{a\})]$, then $f(s, a) = \{a\}$, (3) *Anonymity* iff $(R_i = R'_{\sigma(i)}$ for all $i \in N)$ implies $f(s, A) = f(s', A)$; and (4) *Neutrality* iff [for all $i \in N$, and for all $a, b \in X$, aR_ib iff $\psi(a) R'_i \psi(b)$] implies $f(s', \psi(A)) = \psi[f(s, A)]$.

3. A REVIEW OF SOME ASSUMPTIONS

In this section we comment on three features of Theorem 1: (i) the role of Assumption 2; (ii) the assumption that for all $i \in N$, $S_i = J$; and (iii) the degree of binariness postulated in the theorem.

First, consider the role of Assumption 2 in Theorem 1. A natural question that arises in this context is whether this assumption can be replaced by the following weaker assumption (to be called Assumption 4) without affecting the validity of the result.

ASSUMPTION 4. *For all $i \in N$, and for all non-empty subsets G and G′ of X, $G[\bar{R}_i >| G'$ if for all $x \in G$ and for all $x' \in G'$, $x\bar{P}_i x'$. (Replacing "if" by "only if" in this statement, we get a different assumption to be called Assumption 4′.)*

The following example shows that Assumption 4 cannot replace Assumption 2 in the statement of Theorem 1.

EXAMPLE 3. Let V be any fixed linear ordering over X. Let $S_i = J$ for all $i \in N$. Let the GDF f satisfy the following two conditions:

(1) for every $s \in S$ and for all $x, y \in X$, (xPy iff $[(|\{i \in N \mid xP_i y\}| > |\{i \in N \mid yP_i x\}|)$ or $(|\{i \in N \mid xP_i y\}| = |\{i \in N \mid yP_i x\}|$ and $x \neq y$ and $xVy)]$) and (yRx iff $\sim xPy$); and

(2) for every $s \in S$ and every issue A, if $C(A, R) \neq \emptyset$, then $f(s, A) = C(A, R)$, and if $C(A, R) = \emptyset$, then $f(s, A) = \{x \in A \mid \sim \exists y \in A\colon yP_i x$ for all $i \in N\}$. The reader can check that f satisfies LAV (and hence $\text{LSOV}_{|N|-1}$), and MB, and that f is strictly non-manipulable under Assumption 4′ (which is consistent with Assumption 4).

Theorems 1 and 2 assume tht for all $i \in N$, $S_i = J$ and the possibility of individuals being indifferent between alternatives was used crucially in the proofs. The following theorem shows that if individual orderings (sincere as well as non-sincere) are restricted to be strict, then no longer it is possible to prove that every GDF satisfying MB and $\text{LSOV}_{|N|-1}$ will be manipulable under Assumption 2 when $|X| \geqslant 4$.

THEOREM 3. *If $S_i = J^0$ for all $i \in N$, and if Assumption 3′ is satisfied, then every* GDF *satisfying* LIND, LM, LRS *and* SZ *is strictly non-manipulable.*

Proof. Let f be a GDF satisfying LIND, LM, LRS and SZ. Let $\bar{s}$ be any sincere situation and let (s, L) be a threat to s given some issue A. We show that this leads to a contradiction given that for all $i \in N$, $S_i = J^0$ and that Assumption 3′ holds.

Let $f(\bar{s}, A) = \bar{E}$ and $f(s, A) = E$. Since (s, L) is a threat to $\bar{s}$, $E \neq \bar{E}$. Hence either (i) $\bar{E}$ is a proper subset of E, or (ii) E is a proper subset of $\bar{E}$, or (iii) neither (i) nor (ii) holds.

Suppose (i) holds. By SZ and LRS, $y\bar{P}z$ for all $y \in \bar{E}$ and all $z \in (A - \bar{E})$. By Assumption 3′, $x\bar{P}_L y$ for all $x \in (E - \bar{E})$ and all $y \in \bar{E}$. Since for all $i \in (N - L)$, $R_i = \bar{R}_i$, and for all $x \in (E - \bar{E})$ and all $y \in \bar{E}$, $y\bar{P}x$ and $x\bar{P}_L y$, by LIND and LM it follows that yPx for all $y \in \bar{E}$ and all $x \in (E - \bar{E})$. Since by SZ and LRS, yPz for all $y \in \bar{E}$ and all $z \in (A - E)$, this implies that yPz for all $y \in \bar{E}$ and all $z \in (A - \bar{E})$. This violates SZ since $\bar{E}$ is a proper subset of E.

The proof for case (ii) is exactly similar to the proof for case (i).

Suppose (iii) holds. Then by SZ and LRS, $y\bar{P}x$ for all $y \in (\bar{E} - E)$ and all $x \in (E - \bar{E})$. By Assumption 3′, $x\bar{P}_L y$ for all $x \in (E - \bar{E})$ and all $y \in (\bar{E} - E)$. Since for all $i \in (N - L)$, $R_i = \bar{R}_i$, and for all $x \in (E - \bar{E})$ and all $y \in (\bar{E} - E)$, $y\bar{P}x$ and $x\bar{P}_L y$, by LIND and LM, it follows that yPx for all $y \in (\bar{E} - E)$ and all $x \in (E - \bar{E})$. However, this violates SZ. This completes the proof. ∎

Theorem 3 is a generalization of Theorem 2 of Dutta [5] who proves a similar result with the additional assumption of neutrality. Since Assumption 2′ implies Assumption 3′, the following result follows from Theorem 3.

COROLLARY OF THEOREM 3. *If $S_i = J^0$ for all $i \in N$, and if Assumption 2′ is satisfied, then every* GDF *satisfying* LIND, LM, LRS *and* SZ *is strictly non-manipulable.*

Since Assumption 2′ is consistent with Assumption 2, it is clear that the counterpart of Theorem 1 cannot be proved for the case where for all $i \in N$, $S_i = J^0$.

Note that in Theorem 1, $\text{LSOV}_{|N|-1}$ imposes direct restrictions on outcomes for two-element issues only. Restrictions on outcomes from issues with more than two elements come indirectly only through MB. Thus the hypothesis of the theorem does not impose even Pareto optimality on outcomes from issues with more than two alternatives (though $\mathrm{LSOV}_{|N|-1}$, in the presence of LM, does imply Pareto optimality of outcomes for two-element issues). When $S_i = J^0$, one may explore the possibility of generating impossibility results by imposing direct restrictions on outcomes from issues with more than two elements, in addition to assuming $\mathrm{LSOV}_{|N|-1}$ and MB. The following example shows that when $S_i = J^0$, the properties of Pareto optimality (for every issue), $\mathrm{LSOV}_{|N|-1}$, and MB are not enough to generate manipulability results under Assumption 2′.

Example 4. Let $S_i = J^0$ for all $i \in N$. Let V be a fixed linear ordering over X and let f be a GDF satisfying conditions 1 and 2 in Example 3. Clearly for every issue A and for every situation s, the outcomes under f are Pareto optimal (in terms of s). Also f satisfies $\mathrm{LSOV}_{|N|-1}$ and MB. However, it can be checked that f is strictly non-manipulable under Assumption 2′ (which is consistent with Assumption 2) given that $S_i = J^0$ for all $i \in N$.

When $S_i = J^0$ for all $i \in N$, it may still be possible to generate manipulability results by imposing, in addition to $\mathrm{LSOV}_{|N|-1}$ and MB, properties which are stronger than Pareto optimality and which impose direct restrictions on outcomes from issues with more than two elements. However, this can substantially restrict the scope of the theorem by excluding a large number of GDFs based on pairwise comparisons. Note that, as shown by the following example, even the general property of Pareto optimality may not be fulfilled by plausible decision procedures which proceed by a sequence of pairwise comparisons.

Example 5. Suppose $N = \{1, 2, 3\}$ and $X = \{x_1, x_2, x_3, x_4\}$. Suppose for every $s \in S$, $f(s, X) = X^4(q, s, f)$, where $q(1) = x_1$; $q(2) = x_3$; $q(3) = x_4$; and $q(4) = x_2$ (we are using here notation introduced in Section 2). Suppose for all $s \in S$ and for all $a, b \in X$, aRb iff $|\{i \in N \mid aP_i b\}| \geqslant |\{i \in N \mid bP_i a\}|$. Consider $s \in S$ such that $x_1 P_1 x_2 P_1 x_4 P_1 x_3$; $x_3 P_2 x_1 P_2 x_2 P_2 x_4$; and $x_4 P_3 x_3 P_3 x_1 P_3 x_2$. Then ($x_3 P x_1$ and $x_1 P x_2$ and $x_2 P x_4$ and $x_4 P x_3$) and $f(s, X) = \{x_2\}$, which violates Pareto optimality.

A question of considerable interst is whether in Theorem 1 one can relax the degree of binariness imposed by MB without affecting the negative conclusion. For example, it would be of interest to find out whether the result remains intact if MB is replaced by Sen's [14] condition γ_2 which is much weaker than MB and which only requires that for every issue A and every situation s, if $\{x \in A \mid xPy \text{ for all } y \in (A - \{x\})\} \neq \emptyset$, then $f(s, A) = \{x \in A \mid xPy \text{ for all } y \in (A - \{x\})\}$. We have not been able to resolve this question in either way, and therefore we leave it as an open problem.

REFERENCES

1. K. J. ARROW, "Social Choice and Individual Values," Wiley, New York, 1951; 2nd edition, 1963.
2. J. BLAU AND R. DEB, Social decision functions and the veto, *Econometrica* **45** (1977), 871–879.
3. S. BARBERA, Manipulation of social decision functions, *J. Econ. Theory* **15** (1977), 266–278.
4. S. BARBERA, The manipulation of social choice mechanisms that do not leave "too much" to chance, *Econometrica* **45** (1977), 1573–1588.
5. B. DUTTA, Possibility theorems on strategy-proofness of non-binary social choice functions, mimeograph, 1979.
6. P. GARDENFORS, Manipulation of social choice functions, *J. Econ. Theory* **13** (1976), 217-228.
7. A. GIBBARD, Manipulation of voting schemes: A general result, *Econometrica* **41** (1973), 587–601.
8. A. GIBBARD, Manipulation of schemes that mix voting with chance, *Econometrica* **45** (1977), 665–681.
9. J. S. KELLY, Strategy-proofness and social choice functions without single-valuedness, *Econometrica* **45** (1977), 439–446.
10. P. K. PATTANAIK, "*Strategy and Group Choice*," North-Holland, Amsterdam, 1978.
11. M. A. SATTERTHWAITE, Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *J. Econ. Theory* **10** (1975), 187–217.
12. T. SCHWARTZ, On the possibility of rational policy evaluation, *Theory and Decision* **1** (1970), 89–106.
13. T. SCHWARTZ, Rationality and the myth of the maximum, *Nous* **6** (1972), 97–117.
14. A. K. SEN, Social choice theory: A re-examination, *Econometrica* **45** (1977), 53–89.
15. M. SENGUPTA, Monotonicity, independence of irrelevant alternatives and strategy-proofness of group decision functions, *Rev. Econ. Studies* **47** (1980), 393–408.

JOURNAL OF ECONOMIC THEORY 25, 353–366 (1981)

Wealth and Power in a Collegial Polity*

ROY GARDNER

CEPREMAP, Paris, France,
Iowa State University, Ames, Iowa 50011,
and Northwestern University, Evanston, Illinois 60201

Received July 27, 1979; revised November 30, 1980

I. INTRODUCTION

One of the basic problems in social theory is the relationship between wealth and power. Marxist social theory, for example, argues that political power rests with the wealthy, the only agents who can afford to acquire it. Whatever the merits of such a position, this paper will argue for the relationship in the opposite direction: that wealth follows political power. The framework for doing this is based on the Aumann–Kurz model of taxation [2–4].

In the Aumann–Kurz theory, society takes the form of a continuum of agents, each represented by a utility function and an initial endowment. The endowment (for present purposes a scalar) represents initial wealth. Political power is evenly distributed, and the basic political institution is majority rule. In particular, the majority can exercise the state power to levy taxes. At the same time, each agent retains the right to destroy some or all of his endowment to avoid taxation. That ensuing wealth redistribution game is then analysed in terms of the Shapley value.

The present approach differs from that of Aumann–Kurz in three ways. First, society takes the form of an atom–ocean model. Second, political power is not evenly distributed. Lastly, the basic political institution is not majority rule, but collegial polity. In a collegial polity, a fixed set, called the collegium, has a veto power. The state power can only be exercised by a coalition consisting of the collegium and a specified percentage of the rest of those voting. (This is not the most general possible collegial polity. For further discussion, see Brown [6].) In the present model, the collegium is the

---

* The author wishes to thank F. Breyer, W. Enders and R. Guesnerie for stimulating discussions, and two anonymous referees and an associate editor for extensive commentary. The usual caveat applies. Research support from the National Center for Scientific Research (France) and from the Iowa State University Research Foundation is gratefully acknowledged.

0022-0531/81/060353-14$02.00/0

Copyright © 1981 by Academic Press, Inc
All rights of reproduction in any form reserved.

atom; the rest of those voting, the ocean. The distribution of political power in this collegial polity is measured by the Shapley value (Proposition 1). In general, power is not evenly distributed between the atom and the ocean. The collegial polity tends toward dictatorship when the atom needs very little of the ocean to exercise the state power. However, when the atom needs almost all of the ocean to exercise the state power, the collegial polity tends toward unanimity rule.

Now it is clear why one must continue to allow those being taxed to destroy some or all of their endowment. Otherwise, as the collegium approaches dictatorial political power, it will also approach dictatorial economic power, i.e., confiscation of all endowments (Proposition 3).

The wealth redistribution game then is as follows. Each agent is described by a utility function and an endowment. Redistribution decisions are made by a collegial polity, in which one agent, an atom, has a veto power. Each agent retains the right to destroy some or all of his endowment.

The formal description of society is the subject of the next section. Section III discusses the solution concept, which is the value allocation based on the asymptotic value. Threats are treated in Section IV, and results collected in Section V. There one shows (Proposition 2) that the marginal tax rates of the ocean in the wealth distribution game approach those given by the Aumann–Kurz theory as the collegial polity approaches dictatorship, but approach zero as the collegial polity approaches unanimity rule. In this respect, wealth follows political power. The conclusion considers the problem of generalizing these results.

II. ECONOMIC AND POLITICAL STRUCTURE OF SOCIETY

Society, the set of all agents, is represented by the measurable space $([0,1],\mathscr{B})$, where $\mathscr{B}$ is the collection of Borel subsets of $[0,1]$. Write $T=[0,1]$ and $I=(0,1]$. Let $S\in\mathscr{B}$, and define the functions $X_S:[0,1]\to\{0,1\}$ by

$$\begin{aligned} X_S(t) &= 1 \qquad \text{if} \quad t\in S \\ &= 0 \qquad \text{if} \quad t\in T/S. \end{aligned}$$

The population measure is the measure μ on $(T,\mathscr{B})$ defined by

$$\mu(S)=L(S\cap I)+X_S(0)$$

for each $S\in\mathscr{B}$, where L is Lebesgue measure. In particular,

$$\mu(T)=2, \qquad \mu(\{0\})=1=\mu((0,1]), \tag{1}$$

the point 0 being an atom of the population measure μ.

A coalition $S \in \mathscr{B}$ is winning if it has the power to levy taxes. The complement of a winning coalition is losing. A losing coalition can have its endowment taxed, but cannot levy taxes. A coalition which is neither winning nor losing is indecisive. Indecisive coalitions neither levy taxes, nor are taxed. Let q, $1 < q < 2$, be a fixed parameter. Then,

$$\text{coalition } S \text{ is } \begin{cases} \text{winning} \\ \text{indecisive} \\ \text{losing} \end{cases} \quad \text{as} \quad \begin{cases} q < \mu(S) \\ 2 - q \leqslant \mu(S) \leqslant q \\ \mu(S) < 2 - q \end{cases} . \tag{2}$$

The atom is a member of every winning coalition. Thus, following Brown [6], the collection of winning coalitions is a collegial polity whose collegium is the atom. The complement of the collegium, I, is called the ocean.

It is convenient to introduce the parameter α, which is the measure of the ocean required for a coalition to be winning, given that it contains the atom. The relationship between q and α is simply

$$q = 1 + \alpha \tag{3}$$

As α approaches zero, the collegial polity approaches dictatorship; as α approaches one, unanimity rule.

For each $t \in T$, there is a function u_t on the nonnegative numbers, t's utility, and a nonnegative number $e(t)$, t's initial wealth. Utility is assumed to satisfy the following:

(i) The u_t are increasing, concave, continuously differentiable at positive values of the argument, and continuous at 0.

(ii) The u_t are uniformly bounded, and $u_t(1)$ is uniformly positive, i.e., $\sup_{x,t} u_t(x) < \infty$ and $\inf_t u_t(1) > 0$. (4)

(iii) $u_t(0) = 0$.

(iv) $u_0(x(0)) = x(0)$, and there exists a coalition $S' \subset I$, $\mu(S') > 0$, such that for all $t \in S'$ $u_t(x(t)) = x(t)$.

Assumption (iii) is a normalization, but the rest of these are substantive. Assumption (iv) requires a certain extent of risk-neutrality in the population, in particular that the collegium be risk neutral. The role of assumption (ii), as well as some further technical details on measurability, are spelled out in [3].

As to endowments, the following assumption is made:

$$e(0) = 0, \qquad e(t) \gg 0 \qquad \text{for} \quad t > 0,\ e(T) = 1. \tag{5}$$

The endowment function e is assumed to be measurable, the endowment of coalition $S \in \mathscr{B}$ then being given by

$$e(S) = \int_S e(t)\, d\mu.$$

The assumption $e(T) = 1$ is again a normalization. Since the atom is assumed risk neutral, no generality is lost by setting its endowment at 0.

Finally, an allocation is nonnegative function x on T such that $\int x = \int e$. In light of (5), a positive allocation $x(0)$ to the atom represents taxes collected by the collegium.

III. ASYMPTOTIC VALUE AND VALUE ALLOCATION

A cooperative game is a pair (T, v) consisting of the set of players T and the characteristic function v. This section considers a solution concept for cooperative games, the asymptotic Shapley value, with discussion of v deferred to the next section.

For games with finite T, the Shapley value ϕ is the unique operator on v satisfying the additivity, symmetry, efficiency, and null player axioms [5, Appendix A]. In addition, the Shapley value satisfies the formula

$$\phi v(t) = E[v(S_t^{\mathscr{R}}) - v(S_t^{\mathscr{R}}/\{t\})], \tag{6}$$

where $S_t^{\mathscr{R}}$ is the set of players up to and including t in a random order $\mathscr{R}$ on the set of players, and E is the expectations operator when all orders on T are assigned equal probability. Equation (6) means the Shapley value of a player is the expected value of his marginal contribution to society, given that he is equally likely to occupy any place in a random order. This interpretation will play an important role in the sequel.

The asymptotic Shapley value extends the value concept to certain games with infinite T. The basic idea is as follows. Partition T into a finite number of sets, each set being considered as a player in the corresponding finite game. This game of course has a Shapley value. Now consider a sequence of increasingly fine partitions with a corresponding sequence of Shapley values. If the sequence of Shapley values approaches a limit, and if this limit is independent of the way the sequence is taken, the limit is defined as the asymptotic Shapley value of the infinite game. The asymptotic Shapley value obeys the efficiency axiom and the interpretation (6). For further details, see [5, Chap. 3].

The ability to use the asymptotic Shapley value is provided by the following powerful result due to Neyman [12].

THEOREM. *Let μ be a probability measure on $([0,1], \mathscr{B})$ with finitely many atoms. Suppose $v = f \circ \mu$, where f is a function of bounded variation on $[0,1]$ satisfying $f(0) = 0$ and continuous at 0 and 1. Then v has an asymptotic value.*

All games considered so far have allowed for the transfer of utility. When utility is not transferable, Shapley's device of λ-weights [15] is adopted. Individual utility u_t is scaled by a nonnegative factor $\lambda(t)$, the scaled utilities then being treated as if they were transferable. Thus, the aggregate utility of allocation x for coalition S is

$$\int_S \lambda(t)\, u_t(x(t))\, d\mu.$$

An allocation $x(t)$ is a value allocation if there exist weights $\lambda(t)$ such that

$$\phi v_\lambda(S) = \int_S \lambda(t)\, u_t(x(t))\, d\mu, \qquad \text{all} \quad S \in \mathscr{B}, \tag{7}$$

where

$$v_\lambda(S) = \max_{u_t \in V(S)} \textstyle\int_S \lambda(t)\, u_t(x(t))\, d\mu. \tag{8}$$

At a value allocation, the utility assigned to each coalition by the Shapley value of the λ-weighted game is achieved solely by transfers of wealth; no transfer of utility is required. For an existence result on value allocations, see [13].

IV. THREATS AND THE CHARACTERISTIC FUNCTION

The complicated threat structure inherent in the power to tax and the power to destroy endowments makes the definition of the characteristic function problematic. An important complication is that these threats are variable in nature—the threat to tax all or part of an agent's endowment, for example. To deal with variable threats, Harsanyi's version [7] of Nash's bargaining theory [11] is invoked. Only the highlights of this theory will be discussed here; for a thorough discussion, see [3, Sect. 4].

Let a weighting scheme $\lambda(t)$ be given. If the grand coalition forms, it can clearly achieve the aggregate utility

$$v_\lambda(T) = \max \left\{ \int_T \lambda(t)\, u_t(x(t))\, d\mu : \int_T x = \int_T e \right\}. \tag{9}$$

What a subcoalition S of T can achieve will depend on how susceptible S is to threats a counter-coalitions, in particular, the threats of its complement T/S. For if S and T/S can agree, they can share $v_\lambda(T)$ between themselves; but if they cannot, they must resort to their threats. Let σ and τ be the threat strategies of S and T/S, respectively; $f_\lambda(\sigma, \tau)$ and $g_\lambda(\sigma, \tau)$, the corresponding threat utilities. The optimal threat point (σ^*, τ^*) is the equilibrium point of the two-person zero-sum game with payoff to S of $f_\lambda(\sigma, \tau) - g_\lambda(\sigma, \tau)$. Let

$$w_\lambda(S) = f_\lambda(\sigma^*, \tau^*) - g_\lambda(\sigma^*, \tau^*) = \max_\sigma \min_\tau [f_\lambda(\sigma, \tau) - g_\lambda(\sigma, \tau)]$$

be the value of the optimal threat point to S, with $-w_\lambda(S)$ being the corresponding payoff for T/S. In the Nash bargain, agreement is reached and these optimal threats are not carried out, but the terms of the bargain depend on them:

$$\begin{aligned} v_\lambda(S) &= 1/2[v_\lambda(T) + w_\lambda(S)], \\ v_\lambda(T/S) &= 1/2[v_\lambda(T) - w_\lambda(S)]. \end{aligned} \tag{10}$$

The payoffs (10) define the characteristic function for the wealth redistribution game, once the threat strategies and λ-weights are specified.

V. RESULTS AND EXAMPLES

The first result measures the distribution of power in the collegial polity. The measure of power is the asymptotic value ϕ of the simple game v satisfying

$$\begin{aligned} v(S) &= 1 \qquad \text{if } \mu(S) > q \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{11}$$

PROPOSITION 1. *The asymptotic value of the simple game* (T, v) *satisfying* (11) *is given by*

$$\begin{aligned} \phi v(0) &= 1 - \alpha, \\ \phi v(t) &= \alpha dt, \qquad 0 < t \leqslant 1. \end{aligned} \tag{12}$$

Proof. Consider the finite simple game $[qn; n, 1, 1, \ldots, 1]$ with $n + 1$ players, qn being the quota required for winning, n being the weight of player 0, and the n players with weight 1 each being the finite analogue of the ocean. By a result of Shapiro and Shapley [14], the limiting value of this finite simple game as $n \to \infty$ is

$$\begin{aligned} \phi v(0) &= 2(1-q) = 1-\alpha, \\ \phi v(\{1, 2, \ldots, n\}) &= \alpha. \end{aligned} \tag{13}$$

By Neyman's theorem, the simple game satisfying (11) with T satisfying (1) has an asymptotic value. Hence, (12) follows from (13).

Note that the power measure behaves in an intuitively appealing way. As α approaches 0, the collegial polity approaches dictatorship with the collegium as dictator, but as α approaches 1, the collegium's power erodes. The probabilistic interpretation of (12) based on (7) will also prove useful. Namely, $1-\alpha$ is the probability that the addition of the atom to a randomly drawn coalition will make that coalition winning. In such a case, the atom is said to be *pivotal*. Likewise, the probability that an infinitesimal part of the ocean dt is pivotal is $\alpha\, dt$.

Attention now turns to the characteristic function for the wealth redistribution game. As in [2], define the functions

$$r_\lambda(S) = \max \left\{ \int_S \lambda u(x) \colon x(S) = e(S) \right\} \tag{14}$$

and

$$\begin{aligned} q_\lambda(S) &= r_\lambda(S) && \text{if } \mu(S) \geqslant 2-q \\ &= 0 && \text{if } \mu(S) < 2-q. \end{aligned} \tag{15}$$

r_λ is the von Neumann–Morgenstern characteristic function for the pure exchange situation; q_λ attributes to a coalition its von Neumann-Morgenstern characteristic function value if the coalition is not losing. One can show as in [2] that

$$w_\lambda(S) = q_\lambda(S) - q_\lambda(T/S) \tag{16}$$

and that

$$\phi v_\lambda = \phi q_\lambda. \tag{17}$$

It will prove easier to deal with q_λ in what follows. The dual of game v, denoted $v\#$, is defined by $v\#(S) = v(T) - v(T/S)$. Moreover, $\phi v = \phi v\#$ (see [5, p. 140]). Thus, the distribution of power for the simple game

$$\begin{aligned} v(S) &= 1 && \text{if } \mu(S) \geqslant 2-q \\ &= 0 && \text{otherwise} \end{aligned} \tag{18}$$

is the same as (12), since the games (11) and (18) form a dual pair.

The main result relates the power parameter α to the after-tax distribution of wealth, as given by the value allocation.

PROPOSITION 2. *The value allocation of the wealth redistribution game with power parameter α is given by*

$$x(0) = c(1-\alpha)^2/2, \tag{19}$$

$$\frac{1-\alpha^2}{1-\alpha^2}\,\frac{u_t(x(t))}{u_t'(x(t))} + x(t) = e(t) + \frac{2\alpha(1-\alpha)}{1+\alpha^2}\,c, \tag{20}$$

where c is a positive constant. Moreover,

$$c = (2/(1+2\alpha-\alpha^2))\int_I \frac{u^t(x(t))}{u_t'(x(t))}\,d\mu.$$

Proof. The argument will be fairly informal; a more rigorous argument, along the lines of [3], is possible.

First, some results about r_λ. Let S' be the set of risk-neutral agents in the ocean. If the allocation

$$\begin{aligned} x^1 &= x'(t) && \text{for } t \in S' \\ &= x(t) && \text{for } t \in I/S' \end{aligned}$$

achieves $r_\lambda(I)$, then the allocation

$$\begin{aligned} x^2 &= x(0) && \\ &= x''(t) && \text{for } t \in S' \\ &= x(t) && \text{for } t \in I/S', \end{aligned}$$

where $x(0) + \int_{S'} x''(t) = \int_{S'} x'(t)$ achieves $r_\lambda(T)$, provided both allocations are positive. From the conditions leading to the maximum in (14), and *letting p denote the Lagrangean multiplier corresponding to the wealth constraint, x^1 satisfies*

$$\begin{aligned} p &= \lambda(t), && t \in S', \\ p &= \lambda(t)\, u_t'(x(t)), && t \in I/S', \end{aligned}$$

and $\int x^1 = 1$. At the same time, x^2 satisfies

$$\begin{aligned} p &= \lambda(0) = \lambda(t), && t \in S', \\ p &= \lambda(t)\, u_t'(x(t)), && t \in I/S', \end{aligned}$$

and $\int x^2 = 1$.

Also, it follows from this that $r_\lambda(I) = r_\lambda(T)$. Similarly, all along the diagonal sI, $0 \leqslant s \leqslant 1$, one has

$$r_\lambda(sI) = r_\lambda(\{0\} \cup sI).$$

Moreover, r_λ satisfies the homogeneity property

$$r_\lambda(sI) = s r_\lambda(I).$$

Next, one computes the expected value of the marginal contribution on an agent in a random order.

Consider first the atom 0. Denote by S the set of all agents up to and including 0 in a random order of all agents. The value $\phi q_\lambda(0)$ is the expectation of the contribution $q_\lambda(S) - q_\lambda(S/\{0\})$ of 0 to S. Since the ordering is random and there are many agents, $S/\{0\}$ is a perfect sample of the ocean, $S/\{0\} = sI$, for some s. If $\mu(S) < 2 - q$, i.e., $\mu(S/\{0\}) < 1 - \alpha$, then the contribution of the atom is $r_\lambda(S) = s r_\lambda(T)$. If $\mu(S) \geqslant 2 - q$, i.e., $\mu(S/\{0\}) \geqslant 1 - \alpha$, then the contribution of the atom is zero. Hence, the expected contribution is

$$\int_0^{1-\alpha} s r_\lambda(T)\,ds + \int_{1-\alpha}^{1} 0\,ds = \frac{(1-\alpha)^2}{2} r_\lambda(T). \tag{21}$$

Consider next an infinitesimal agent dt of the ocean, S being a random order as before. With probability 1/2, the atom is before dt in the ordering. In this case, dt's contribution to S is

$$r_\lambda(S) - r_\lambda(S/dt).$$

In view of homogeneity, however, this contribution is the same as dt's contribution to T. The latter contribution is shown in [1] to be

$$(\lambda(t)\, u_t(x(t)) - p(x(t) - e(t)))\,\mu(dt). \tag{22}$$

With probability 1/2, the atom is after dt in the ordering. The conditional probability then is $(1-\alpha^2)$ that the coalition S is losing, in which case dt contributes nothing, and α^2 that the coalition S is indecisive, in which case dt's contribution is again (22). Finally, with probability $\alpha\mu(dt)$, dt is pivotal. In this case, dt's contribution is $r_\lambda((1-\alpha)I)$, since $\mu(S) = 1 - \alpha$ for dt to be pivotal. Summing up, dt contributes nothing with probability $(1-\alpha^2)/2$, $(1-\alpha)\, r_\lambda(I)$ with probability $\alpha\mu(dt)$, and (22) with probability $(1+\alpha^2)/2$, so his expected marginal contribution is

$$(\phi q_\lambda)(dt) = \left[\frac{1+\alpha^2}{2}(\lambda(t)\, u_t(x(t)) - p(x(t) - e(t))) + \alpha(1-\alpha)\, r_\lambda(T)\right]\mu(dt). \tag{23}$$

If x is a value allocation, then $u(x)$ is a value, i.e.,

$$\lambda(t)\,u_t(x(t))\mu(dt)=\phi v_\lambda(dt) \qquad \text{for } dt \text{ in the ocean,} \tag{24}$$

$$\lambda(0)\,u_0(x(0))\,\mu(0)=v_\lambda(0) \qquad \text{for the atom.} \tag{25}$$

Combining this with the help of (17), (21), and (23),

$$\lambda(0)\,x(0)=((1-\alpha)^2/2)\,r_\lambda(T), \tag{26}$$

$$\lambda(t)\,u_t(x(t))=\frac{1+\alpha^2}{2}\,(\lambda(t)\,u_t(x(t))-p(x(t)-e(t)))+\alpha(1-\alpha)\,r_\lambda(T), \tag{27}$$

whence

$$x(0)=((1-\alpha)^2/2)\,r_\lambda(T)/p, \tag{28}$$

$$\frac{1-\alpha^2}{1+\alpha^2}\,\frac{u_t(x(t))}{u_t'(x(t))}+x(t)=e(t)+\frac{2\alpha(1-\alpha)\,r_\lambda(T)}{(1+\alpha^2)\,p}. \tag{29}$$

It follows immediately that $x(t)>0$, so that $p=\lambda(t)\,u_t'(x(t))$. Substituting, and setting $c=r_\lambda(T)/p$, one has (19), (20).

Since $x(t)$ is an allocation, $\int_T x(t)=\int_T e(t)$. It then follows from (19), (20) that

$$c=(2/(1+2\alpha-\alpha^2))\int_t\frac{u_t(x(t))}{u_t'(x(t))}.$$

The influence of political power on the final distribution of wealth is evident from (19) and (20). First, $x(0)>0$ for all α; the collegium is always a net tax collector. However, $x(0)$ approaches zero as α approaches one and the collegium's power evaporates. Second, there is a direct relationship between the marginal tax rates paid by members of the ocean and the distribution of power. The total tax burden of dt is $e(t)-x(t)$, and his marginal tax rate is the drivative of this expression with respect to (t). A routine calculation shows that

$$\frac{d}{de(t)}(e(t)-x(t))=1-\frac{1+\alpha^2}{2-(1+\alpha^2)\dfrac{u_t u_t''}{(u_t')^2}}. \tag{30}$$

As α approaches zero in (30), the members of the ocean pay Aumann–Kurz marginal tax rates [2, Sect 5], while as α approaches one, the marginal tax rate vanishes. The minimum marginal tax rate, $(1-\alpha^2)/2$, is paid by a risk neutral agent; the more risk-averse an agent, the greater his marginal tax rate. The overall tendence is for marginal tax rates to fall as the ocean becomes more powerful. Nevertheless, an agents's total tax burden rise or fall with α, as the following examples show.

Example 1 (Linear utility). In this example, $u_t(x(t)) = x(t)$ for all agents. For agents in the ocean,

$$\frac{1-\alpha^2}{1+\alpha^2}\frac{u}{u'} + x = x(2/(1+\alpha^2)),$$

so that from (20)

$$x = \frac{1+\alpha^2}{2}e + \alpha(1-\alpha)c. \tag{31}$$

Integrating, $c = 1$. Thus, the value allocation is

$$x(0) = (1-\alpha)^2/2,$$

$$x(t) = \frac{1+\alpha^2}{2}e(t) + \alpha(1-\alpha).$$

The tax collected by the collegium falls monotonically with decreases in the collegium's power. The total tax burden for an agent in the ocean is

$$\frac{1-\alpha^2}{2}e(t) - \alpha(1-\alpha).$$

This burden is always positive as long as $e(t) > 1$. In view of the normalization (5), this means that above-average wealthholders always pay a tax. However, below-average wealthholders, at some values of α, are actually subsidized (negative total tax burden). For these agents, total burden as a function of α is positive at $\alpha = 0$ and zero at $\alpha = 1$. It is negative all along the interval $e(t)/(2-e(t)) \leqslant \alpha < 1$. Differentiation shows that the total tax burden of above-average wealthholders decreases as their political power increases. For a below-average wealthholders, this is only true up to the point where $\alpha = 1/(2-e(t))$, the point where the subsidy is largest. After this point, the subsidy falls as political power rises, until it vanishes at $\alpha = 1$. Clearly, the increasing political power of the ocean makes its richer members less susceptible to the redistributive pressures of those poorer, at least after the maximum subsidy point.

Example 2. In this example, one half of the agents in the ocean are risk-neutral, but the rest are risk-averse with utility $u_t(x(t)) = x(t)^\gamma$. The risk-neutral agents continue to satisfy (31). For the risk-averse agents,

$$\frac{1-\alpha^2}{1+\alpha^2}\frac{u}{u'} + x = x\left(\frac{1-\alpha^2}{1+\alpha^2}\frac{1}{\gamma} + 1\right),$$

hence,

$$x = \frac{\gamma(1+\alpha^2)}{1-\alpha^2+\gamma(1+\alpha^2)}e + \frac{2\alpha(1-\alpha)\gamma}{1-\alpha^2+\gamma(1+\alpha^2)}c. \tag{32}$$

Assuming the risk-neutral and risk-averse sectors share pretax wealth equally, one obtains by integration

$$c = \frac{[\gamma(1+\alpha^2)+3-\alpha^2]}{2[1-\alpha^2+\gamma(1+\alpha^2)]}.$$

Since $c > 1$ for all α, all risk-neutral agents receive a larger after-tax allocation here than in the previous example. Their gain comes at the expense of the risk-averse sector of the ocean. This recalls a feature noted by Kurz [9, 10], that a risk-averse agent does better to conceal his risk-aversion from others. Even though $dc/d\alpha > 0$, it is still the case that $dx(0)/d\alpha < 0$; erosion of the collegium's political power lessens its tax collection. Finally, among the risk-averse agents, only those whose endowments satisfy

$$e(t) < \frac{\alpha\gamma(\gamma(1+\alpha^2)+3-\alpha^2)}{(1+\alpha)(1-\alpha^2+\gamma(1+\alpha^2))}$$

qualify for a subsidy. For all risk-averse agents, however, the marginal tax rate is a decreasing function of their political power.

To assess the role of the threat by losing coalitions to destroy their endowment, one can consider the wealth redistribution game where this threat is not made. The characteristic function is given directly by

$$\begin{aligned} v_\lambda(S) &= r_\lambda(T) & & \text{winning} \\ &= r_\lambda(S) & \text{if } S \text{ is} \quad & \text{indecisive} \\ &= 0 & & \text{losing} \end{aligned} \tag{33}$$

since a winning coalition can collect all endowments as taxes.

Proposition 3. *The value allocation of the wealth redistribution game* (33) *with given power marameter* α *satisfies*

$$x(0) = (1-\alpha)^2 c, \tag{34}$$

$$\frac{(1-\alpha^2)}{2}\frac{u_t(x(t))}{u_t'(x(t))} + x(t) = e(t) + 2(1-\alpha)c/\alpha, \tag{35}$$

where

$$c = (1/(\alpha(2-\alpha))) \int_I \frac{u_t(x(t))}{u_t'(x(t))}.$$

The proof is like that of Proposition 2. Removing the threat of endowment destruction gives the collegium both the economic and the political power of a dictator as α approaches zero. By the same token, for fixed α (35) implies higher marginal tax rates than (20) for the ocean. In the specific case of Example 1, (34) represents a doubling of the collegium's tax collection, while (35) represents a doubling of the marginal tax rates for the ocean. Paradoxically, an agent in the ocean may do better without this threat. An agent whose endowment satisfies

$$e(t) < 2\alpha/(1+\alpha)$$

does better when endowments cannot be destroyed. This again is a relative wealth effect. Above-average wealthholders always do better when they can destroy their wealth rather than have it confiscated. Whether this is true for below-average wealthholders depends on the distribution of power.

VI. CONCLUSION

The present paper has characterized the value allocation of a wealth redistribution game in which there is an atom with a veto power in the distribution of agents. The crucial assumption in the analysis has been that the atom, as well as a positive measure of the ocean, be risk-neutral. Without this assumption, a simple characterization of the value allocation, like that of (19) and (20), is no longer possible. The marginal contribution of an oceanic agent will depend on the presence or absence of the atom. In a few cases, it is still easy to characterize a value allocation. For instance, a risk-averse atom fares worse than this risk-neutral counterpart against a risk neutral ocean. Nevertheless, characterizing the value allocation when all agents are risk-averse remains an open question.

REFERENCES

1. R. J. AUMANN, Values of markets with a continuum of agents, *Econometrica* **43** (1975), 611–646.
2. R. J. AUMANN AND M. KURZ, Power and taxes, *Econometrica* **45** (1977), 1137–1161.
3. R. J. AUMANN AND M. KURZ, Power and taxes in a multi-commodity economy, *Israel J. Math.* **27** (1977), 185–234.

4. R. J. Aumann and M. Kurz, Power and taxes in a multi-commodity economy (updated), *J. Public Econ.* **9** (1978), 189–161.
5. R. J. Aumann and L. S. Shapley, "Values of Non-Atomic Games," Princeton Univ. Press, Princeton, N.J., 1974.
6. D. Brown, Aggregation of preferences. *Quart. J. Econ.* **89** (1975), 456–469.
7. J. C. Harsanyi, A bargaining model for the cooperative *n*-person game, *in* "Contributions to the Theory of Games IV" (Luce and Tucker, Eds.), 325–356, Princeton Univ. Press, Princeton, N.J., 1974.
8. M. Kurz, Distortion of preferences, income distribution, and the case for a linear income tax, *J. Econ. Theory* **14** (1977), 291–298.
9. M. Kurz, Income distribution and distortion of preferences: the *l*-commodity case, *J. Econ. Theory* **22** (1980), 99–106.
10. J. F. Nash, The bargaining problem, *Econometrica* **18** (1950), 155–162.
11. J. F. Nash, Two person cooperative games, *Econometrica* **21** (1953), 128–140.
12. A. Neyman, "Asymptotic Values of Mixed Games," Technical Report No. 289, Stanford University, 1979.
13. A. Neyman, "Values of Non-Transferable Utility Games with a Continuum of Players," Technical Report No. 351, Cornell University, 1977.
14. N. Z. Shapiro and L. S. Shapley, Values of large games I: A limit theorem, *Math. Operations Res.* **3** (1978), 1–9.
15. L. S. Shapley, Utility comparisons and the theory of games, *in* "La Décision" (Guilbaud, Ed.), 251–263, Editions du CNRS, Paris, 1969.

JOURNAL OF ECONOMIC THEORY 25, 367–379 (1981)

On Pareto-Efficiency and the No-Envy Concept of Equity

KOTARO SUZUMURA*

Institute of Economic Research, Kyoto University,
Yoshidahonmachi, Sakyo-ku, Kyoto, Japan

Received January 24, 1980; revised October 30, 1980

1. INTRODUCTION

In dividing some fixed amount of resources among a fixed number of individuals, the exclusive reliance on the Pareto-efficiency criterion will be of little help, since *too many* feasible divisions will be Pareto-efficient under the "standard" environmental conditions. The introduction of an additional criterion of equity to the effect that a division is *equitable* if and only if no individual envies the position of another individual when the specified division is implemented will substantially narrow down the range of eligible divisions.[1] But the joint use of the Pareto-efficiency and the no-envy concept of equity encounters difficulty when the amount to be divided depends upon the contribution made by individuals, among whom ability differential prevails, since there may then be *no* eligible division even under the "standard" environmental conditions.[2] One possible response to this dilemma would be to observe that, from the viewpoint of moral philosophy, it is not altogether clear whether a concept of equity based on envy can be ethically relevant in the first place and to wash one's hands of the business. The second and arguably more "fruitful" response would be to propose a modified definition of equity which, coupled with the Pareto-efficiency, provides us with an alternative definition of eligibility. In the literature, we have abundantly many proposals to this effect: *wealth-fairness* (Varian [20,

---

* This paper was written while I was visiting Stanford University in 1979–1980. I am deeply indebted to Professors Kenneth Arrow, Peter Coughlin, Peter Hammond, Steve Goldman, and David Starrett for their helpful comments on an earlier draft of this paper. *Thanks are also due to the anonymous referee of this Journal, whose incisive comments and remarks were instrumental in preparing the present version. Needless to say, I am solely* responsible for any defects which may still remain.

[1] This concept of equity is due originally to Foley [7, Sect. IV]. See also Pazner and Schmeidler [10] and Varian [20].

[2] This dilemma was first exposed by Pazner and Schmeidler [10].

0022-0531/81/060367-13$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

21]), *income-fairness* (Varian [20] and Pazner [9]), ***balanced-with-respect-to-envy-justice*** (Daniel [5]), ***egalitarian-equivalence*** (Pazner and Schmeidler [11]) and ***fairness-equivalence*** (Pazner [9]).[3] The third response would be to imbed the proposed concept of fairness as equity-cum-efficiency into the conceptual framework of social choice theory with a view of evaluating how the identified difficulty would be located in the perennial enigma of designing "satisfactory" collective choice mechanisms. The present paper is an attempt in this third category of exercises which are related to equity, envy and efficiency. We work with a model of social choice along the line of Arrow [3, Chap. VIII, Sect. IV.4], Sen [16, Chap. 9*] and Suppes [18], in which the aggregation problem is to map each profile of individuals' extended orderings of $X \times N$ into a choice function on the family of subsets of X, where X and N denote the set of social states and the set of individuals, respectively. In this set-up, the viability of Foley's [7] original definition of fairness will be critically examined.

2. FAIRNESS-AS-NO-ENVY-CUM-EFFICIENCY AND A GENERALIZED COLLECTIVE CHOICE RULE

2.1. As Varian [21, p. 240] has aptly observed, "[t]he theory of fairness ... is founded in the notion of 'extended sympathy' and in the idea of 'symmetry' in the treatment of agents. ... In effect, we are asking each agent to put himself in the position of each other agents to determine if that is a better or a worse position than the one he is now in." To formalize this foundation of the theory of fairness, let X and $N = \{1, 2, \ldots, n\}$ $(2 \leq n < +\infty)$ stand, respectively, for the set of all conceivable social states and the set of individuals, a social state being the complete description of the relevant aspects of the world. For each $i \in N$, we describe his/her views on the society by an extended preference ordering $\tilde{R}_i$ on the Cartesian product $X \times N$,

$$((x, j), (y, k)) \in \tilde{R}_i$$

denoting the fact that being in the position of individual j in the social state x is at least as good as being in the position of individual k in the social state y according to i's view.[4] A list of extended preference orderings, one ordering for each individual, will be called a *profile* and alternative profiles will be indexed by $\alpha, \beta, \ldots$ like $\alpha = (\tilde{R}_1^\alpha, \tilde{R}_2^\alpha, \ldots, \tilde{R}_n^\alpha)$, $\beta = (\tilde{R}_1^\beta, \tilde{R}_2^\beta, \ldots, \tilde{R}_n^\beta)$ and so on.

---

[3] These proposals were succinctly surveyed and critically evaluated by Pazner [9] and Sen [17, Sect. 5].

[4] *Being an ordering, $\tilde{R}_i$ is connected, reflexive and transitive binary relation on $X \times N$.*

The set of all logically possible profiles will be denoted by $\mathscr{A}$, while the set of all non-empty finite subsets of X will be written as $\mathscr{S}$, each and every $S \in \mathscr{S}$ being construed to represent a set of available states under the specified environmental conditions.

2.2. Take a profile $a = (\tilde{R}_1^a, \tilde{R}_2^a, ..., \tilde{R}_n^a) \in \mathscr{A}$ and an $S \in \mathscr{S}$ and fix them for the time being. For each individual $i \in N$, let i's *subjective preference ordering* R_i^a be defined by

$$R_i^a = \{(x, y) \in X \times X \mid ((x, i), (y, i)) \in \tilde{R}_i^a\}, \tag{1}$$

in terms of which the *a-Paretian quasi-ordering* R_N^a will be defined by

$$R_N^a = \bigcap_{i \in N} R_i^a. \tag{2}$$

The set of all a-Pareto-efficient states in S will then be denoted as follows:

$$E_f^a(S) = \{x \in S \mid \sim[\exists y \in S: (y, x) \in P(R_N^a)]\}, \tag{3}$$

where and hereafter $P(\cdot)$ will denote an operator giving the asymmetric part of the binary relation in the parenthesis and $\sim$ denotes logical negation.

Following Foley's [7] classical definition, we say that individual $i \in N$ envies individual $j \in N$ at $x \in X$ when the profile a prevails if and only if $((x, j), (x, i)) \in P(\tilde{R}_i^a)$ holds true. We say that x is *a-equitable* if and only if nobody envies other individuals at x when the profile a prevails. The set of all a-equitable states in $S \in \mathscr{S}$ may be written as follows:

$$E_q^a(S) = \{x \in S \mid \forall i, j \in N: ((x, i), (x, j)) \in \tilde{R}_i^a\}. \tag{4}$$

If a state in S is simultaneously a-Pareto-efficient and a-equitable, it is said to be *a-fair* in S. The set of all a-fair states in S, to be called the a-fair set in S, will be denoted by

$$F^a(S) = E_f^a(S) \cap E_q^a(S). \tag{5}$$

2.3. The first point to be clarified about the a-fair set in S is that it may well be empty. Indeed, it is even possible that $E_q^a(X) = \varnothing$ for some $a \in \mathscr{A}$, namely, there may exist no a-equitable state wheresoever. Even if a-Pareto-efficiency and a-equity are *individually* self-consistent (in the sense that they may respectively be satisfied), they may well be *jointly* incompatible as is easily exemplified as follows.

EXAMPLE 1. Let $X = \{x, y\}$ and $N = \{1, 2\}$. Let a profile α be specified by

$$\tilde{R}_1^\alpha: (y, 2), (y, 1), (x, 1), (x, 2),$$
$$\tilde{R}_2^\alpha: (y, 2), (y, 1), (x, 2), (x, 1).^5$$

Clearly $E_q^\alpha(\{x, y\}) = \{x\}$ and $P(R_N^\alpha) = \{(y, x)\}$, so that we have $F^\alpha(\{x, y\}) = \varnothing$. ∎

Simple though this example is, it may serve us well to expose several important features of the α-fairness concept. As an auxiliary step, let R_{ij}^α be defined by

$$R_{ij}^\alpha = \{(x, y) \in X \times X \mid ((x, j), (y, j)) \in \tilde{R}_i^\alpha\}. \qquad (6)$$

Clearly it is true that $R_i^\alpha = R_{ii}^\alpha$ for all $\alpha \in \mathscr{A}$ and all $i \in N$. Notice that $(x, y) \in R_i^\alpha$ means that individual i thinks that x is no worse for him than y, while $(x, y) \in R_{ij}^\alpha$ means that i thinks that it is no worse for j to be in x rather than in y. With this interpretation of R_i^α and R_{ij}^α, it seems fairly natural in the context where we talk about the welfare judgements based on the extended sympathy that we require the fulfillment of the following axiom, which is due to Sen [16, p. 156].[6]

AXIOM OF IDENTITY.

$$\forall i, j \in N: R_{ij}^\alpha = R_j^\alpha.$$

It is well recognized that the lack of the sympathetic acceptance of other's subjective preferences, namely, the invalidity of the axiom of identity, causes many logical difficulty in the exercise of aggregating profiles of extended preference orderings.[7] Notice, however, that the profile specified in the Example 1 does satisfy the axiom of identity, so that the problem identified by this example emerges even if the sympathetic identification prevails among individuals.

---

[5] Preference orderings will be written horizontally with more preferred state–individual combination to the left of less preferred, indifferent combinations (if any) being put together by square brackets.

[6] What the axiom of identity requires is that "placing oneself in the position of the other should involve not merely having the latter's objective circumstances but also identifying oneself with the other in terms of his subjective features" (Sen [16, pp. 149–150]). It is debatable, however, if indeed we need literal transformation of subjective features so as to comply with the requirement of the axiom. On this and related points, the interested readers are referred to Suzumura [19].

[7] See, for example, Sen [16, pp. 149–150] and Suzumura [19].

Our second remark on the α-fairness concept concerns with the contrast between the justice concept thereby implied and the traditional rival justice concepts, i.e., the ***Rawlsian leximin justice*** and the ***Benthamite utilitarian justice***. Notice that in the situation specified by the Example 1 *both* a Rawlsian and a Benthamite would assert that y is more just than x—assuming for the sake of gaining comparability that 1 and 2 have interpersonally fully comparable cardinal representation of $\tilde{R}_1^\alpha$ and $\tilde{R}_2^\alpha$, respectively, while $E_q^\alpha(\{x, y\}) = \{x\}$ would force one to say that x is more equitable than y. The ethical appeal of the fairness-as-no-envy approach seems to be rather fragile indeed.

3. ON THE POSSIBILITY OF FOLEY–FAIR COLLECTIVE CHOICE RULES

3.1. Our problem is to design a "fair" generalized collective choice rule Ψ, GCCR for short, which amalgamates each profile $a \in \mathscr{A}$ of extended preference orderings into a social choice function $C^a = \Psi(a)$ on $\mathscr{S}$ such that, for each set $S \in \mathscr{S}$ of available social states, $C^a(S)$ denotes the non-empty set of chosen states reflecting a "fair" amalgamation of a we have started from. In view of the possible non-existence of a-fair states for some $a \in \mathscr{A}$, care should be taken with the sense in which we mean a GCCR to be "fair." One sense which naturally suggests itself is to require that Ψ satisfies the following condition.

FAIRNESS EXTENSION (FE). *For each admissible profile* a, $C^a = \Psi(a)$ *satisfies* $F^a(S) = C^a(S)$ *whenever* $S \in \mathscr{S}$ *is such that* $F^a(S) \neq \varnothing$.

3.2. Collective choice is a repeated exercise in changing environments and one naturally feels that successive choices made should satisfy some "reasonable" choice-consistency condition. A choice-consistency condition which is deeply rooted in the Arrovian social choice theory is that of collective full-rationality (FR), which requires that we may construe a choice function C to describe a behaviour of optimizing fully consistent collective preference relation. Formally C satisfies the condition FR if and only if there exists a preference ordering R_C on X satisfying

$$\forall S \in \mathscr{S}: C(S) = \{x \in S \mid \forall y \in S: (x, y) \in R_C\}.$$

It was shown by Arrow [2] that a choice function C on $\mathscr{S}$ satisfies this condition FR if and only if C satisfies the following axiom.

ARROW'S AXIOM (AA).

$$\forall S_1, S_2 \in \mathscr{S}: S_1 \subset S_2 \rightarrow [S_1 \cap C(S_2) = \varnothing \vee S_1 \cap C(S_2) = C(S_1)].$$

The following two axioms, which were found useful in various social and individual choice contexts, provide a natural decomposition of Arrow's axiom.

CHERNOFF'S AXIOM (CA).

$$\forall S_1, S_2 \in \mathscr{S}: S_1 \subset S_2 \rightarrow [S_1 \cap C(S_2) = \varnothing \vee S_1 \cap C(S_2) \subset C(S_1)].$$

DUAL-CHERNOFF AXIOM (DCA).

$$\forall S_1, S_2 \in \mathscr{S}: S_1 \subset S_2 \rightarrow [S_1 \cap C(S_2) = \varnothing \vee S_1 \cap C(S_2) \supset C(S_1)].$$

Another class of important choice-consistency conditions is that of path-independence, due originally to Arrow [3, Chap. VIII, Sect. V] and Plott [12], and various variants thereof. They essentially require that the choice from a set should be independent of the path to be followed *en route* in search for the global choice.

PATH-INDEPENDENCE (PI).

$$\forall S_1, S_2 \in \mathscr{S}: C(S_1 \cup S_2) = C(C(S_1) \cup S_2).$$

WEAK PATH-INDEPENDENCE α (WPI(α)).

$$\forall S_1, S_2 \in \mathscr{S}: C(S_1 \cup S_2) \subset C(C(S_1) \cup S_2).$$

WEAK PATH-INDEPENDENCE β (WPI(β)).

$$\forall S_1, S_2 \in \mathscr{S}: C(S_1 \cup S_2) \supset C(C(S_1) \cup S_2).$$

Finally we introduce two very weak choice-consistency conditions which still have bites.

SUPERSET AXIOM (SUA).

$$\forall S_1, S_2 \in \mathscr{S}: [S_1 \subset S_2 \,\&\, C(S_2) \subset C(S_1)] \rightarrow C(S_1) = C(S_2).$$

STABILITY AXIOM (ST).

$$\forall S \in \mathscr{S}: C(C(S)) = C(S).$$

To facilitate recollection and later reference, we summarize the logical relationship which holds true among these choice-consistency axioms in the following theorem, where an arrow indicates a logical implication which cannot be reversed in general, while the axioms in square brackets are equivalent to the axiom above them.

THEOREM 1.

$$\begin{tikzcd}
\mathrm{CA} \arrow[d, leftrightarrow] & \begin{array}{c}\mathrm{FR} \\ {[\mathrm{AA} = \mathrm{CA} + \mathrm{DCA}]}\end{array} \arrow[l] \arrow[r] \arrow[d] & \mathrm{DCA} \arrow[d] \\
\mathrm{WPI}(\alpha) \arrow[d] & \begin{array}{c}\mathrm{PI} \\ {[\mathrm{WPI}(\alpha) + \mathrm{WPI}(\beta)]}\end{array} \arrow[l] \arrow[r] & \mathrm{WPI}(\beta) \arrow[d] \\
\mathrm{ST} & & \mathrm{SUA}
\end{tikzcd}$$

Proof. Most of the assertions being either immediate results of the definitions or already established in Blair *et al.* [4], Ferejohn and Grether [6] and Plott [12], we have only to prove that (a) DCA *implies* WPI(β), and (b) WPI(β) *implies* SUA.

(a) Assume that C satisfies the condition DCA on $\mathscr{S}$ and take any $S_1, S_2 \in \mathscr{S}$. Since $C(S_1) \cup S_2 \subset S_1 \cup S_2$ is obviously true, we may invoke DCA to assert that either

$$C(S_1 \cup S_2) \cap [C(S_1) \cup S_2] = \varnothing, \tag{7}$$

or

$$C(S_1 \cup S_2) \cap [C(S_1) \cup S_2] \supset C(C(S_1) \cup S_2) \tag{8}$$

is true. If (8) is indeed the case, we have $C(S_1 \cup S_2) \supset C(C(S_1) \cup S_2)$ and we are home. Assume therefore that (7) is true. Since $S_1 \subset S_1 \cup S_2$ is true, the second use of DCA yields either

$$C(S_1 \cup S_2) \cap S_1 = \varnothing, \tag{9}$$

or

$$C(S_1 \cup S_2) \cap S_1 \supset C(S_1). \tag{10}$$

It follows from (7) that $C(S_1 \cup S_2) \cap C(S_1) = C(S_1 \cup S_2) \cap S_2 = \emptyset$, so that we obtain

$$C(S_1 \cup S_2) \subset S_1 \setminus C(S_1), \tag{11}$$

which negates the validity of (9). Therefore (10) must be true, which however contradicts (11). This concludes our proof of (a).

(b) Assume that C satisfies WPI(β) and let $S_1, S_2 \in \mathscr{S}$ be such that $S_1 \subset S_2$ and $C(S_2) \subset C(S_1)$. Thanks to WPI(β) we then have

$$C(S_2) = C(S_1 \cup S_2) \supset C(C(S_2) \cup S_1) = C(S_1),$$

which, coupled with $C(S_2) \subset C(S_1)$, yields $C(S_1) = C(S_2)$, as desired. ∎

Notice that these choice-consistency axioms are properties of a choice function but they may be regarded as properties of a generalized collective choice rule which generates a choice function having the designated properties. With this understanding in mind we will talk about choice-consistency of a GCCR in the following.

3.3. Let us introduce a rather mild unrestricted domain condition on Ψ, which requires that the class of admissible profiles be rich enough to the following extent.

UNRESTRICTED DOMAIN UNDER THE AXIOM OF IDENTITY (UID). *The domain of Ψ consists of all logically possible profiles satisfying the axiom of identity.*

We are now ready to present our first negative theorem on the "fair" GCCRs.

THEOREM 2. *Suppose that there exist at least three social states. Then there exists no GCCR which satisfies UID (Unrestricted Domain under the Axiom of Identity), FE (Fairness Extension), and SUA (Superset Axiom of Choice-Consistency).*

Proof. Take three distinct social states x, y, and z and let $S_1 = \{x, y\}$ and $S_2 = \{x, y, z\}$. Let a profile $\alpha = (\tilde{R}_1^\alpha, \tilde{R}_2^\alpha, \ldots, \tilde{R}_n^\alpha)$ be such that

$$\tilde{R}_1^\alpha(S_2 \times \{1, 2\}) \colon (x, 1), (z, 2), (z, 1), (y, 1), (y, 2), (x, 2),$$

$$\tilde{R}_2^\alpha(S_2 \times \{1, 2\}) \colon (z, 2), (y, 2), (x, 2), (x, 1), (z, 1), (y, 1),$$

$$\forall i \in N \setminus \{1, 2\} \colon \tilde{R}_i^\alpha(S_2 \times \{1, 2\}) = \tilde{R}_1^\alpha(S_2 \times \{1, 2\}),$$

where $\tilde{R}^a_j(S_2 \times \{1, 2\})$ denotes the restriction of $\tilde{R}^a_j$ on $S_2 \times \{1, 2\}$ for all $j \in N$, and that

$$\forall (v,j) \in (X \times N)\setminus(S_2 \times \{1, 2\}): \begin{cases} ((x, 2), (v,j)) \in P(\tilde{R}^a_1), \\ ((y, 1), (v,j)) \in P(\tilde{R}^a_2), \\ \forall i \in N\setminus\{1, 2\}: ((v,j), (x, 1)) \in P(\tilde{R}^a_i), \end{cases}$$

$$\forall i \in N, \forall (v^1,j^1), (v^2,j^2) \in (X \times N)\setminus(S_2 \times \{1, 2\}):$$
$$((v^1,j^1), (v^2,j^2)) \in I(\tilde{R}^a_i).$$

It is clearly the case that this profile satisfies the axiom of identity. Notice that $E^a_q(S_2) = \{x, y\}$ and $P(R^a_N) \cap (S_2 \times S_2) = \{(z, y)\}$. Therefore $C^a(S_1) = F^a(S_1) = S_1$ and $C^a(S_2) = F^a(S_2) = \{x\}$, where use is made of the condition FE. We then obtain $S_1 \subset S_2$, $C^a(S_2) \subset C^a(S_1)$ and $C^a(S_1) \neq C^a(S_2)$. Therefore a GCCR satisfying UID and FE cannot possibly satisfy SUA. ∎

3.4. The condition FE demands that the fair state and only the fair state should be chosen when one exists. What if there exists no fair state? To consider this situation, let us define an auxiliary binary relation R^a_F on X for each $a \in \mathscr{A}$ by

$$R^a_F = \{(x, y) \in X \times X \mid x \in E^a_q(X) \;\&\; y \notin E^a_q(X)\}. \tag{12}$$

Now if we really care about the appeal of equity-as-no-envy as well as Pareto-efficiency, the following requirement may seem to be reasonable, which says basically that a state which is either "more equitable" or "more efficient" than a state which is chosen should itself be among chosen states.

FAIRNESS INCLUSION (FI). *If $a \in \mathscr{A}$ and $S \in \mathscr{S}$ are such that $F^a(S) = \varnothing$, then*

(a) $[x \in S, (x, y) \in R^a_E \;\&\; y \in C^a(S)] \rightarrow x \in C^a(S)$, *and*
(b) $[x \in S, (x, y) \in P(R^a_N) \;\&\; y \in C^a(S)] \rightarrow x \in C^a(S)$, *where* $C^a = \Psi(a)$.

We also introduce a variant of the Pareto unanimity requirement on a GCCR, but we need careful step forward in this slippery area. To require the exclusion of a state y from a choice set for a binary choice environment $\{x, y\}$ just because x happens to Pareto-dominate y would be grossly inappropriate in the context where we care about equity and the like, since doing so means to empower the Pareto-dominance relation to always outweigh the equity consideration in the binary choice situation. But this lopsided sanctification of the Pareto dominance quite simply contradicts the emphasis put on the equity consideration in the fairness approach. This argument, if accepted, would lead us to the following conditional variant of the Pareto rule.

CONDITIONAL BINARY EXCLUSION PARETO (CBEP). *If an admissible profile* α *and* $x, y \in X$ *are such that* $E_q^\alpha(\{x, y\}) = \varnothing$ *and* $(x, y) \in \bigcap_{i \in N} P(R_i^\alpha)$, *then* $\{x\} = C^\alpha(\{x, y\})$, *where* $C^\alpha = \Psi(\alpha)$.

What emerges out of these mild-looking conditions on a GCCR is another impossibility theorem, which reads as follows.

THEOREM 3. *Suppose that there exist at least three social states. Then there exists no GCCR which satisfies UID (Unrestricted Domain under the Axiom of Identity), FI (Fairness Inclusion), CBEP (Conditional Binary Exclusion Pareto), and CA (Chernoff's Axiom of Choice-Consistency).*

Proof. Take three distinct social states x, y, and z and let $S_1 = \{x, y\}$ and $S_2 = \{x, y, z\}$. Let a profile $\alpha = (\tilde{R}_1^\alpha, \tilde{R}_2^\alpha, \ldots, \tilde{R}_n^\alpha)$ be such that

$$R_1^\alpha(S_2 \times \{1, 2\})\colon (x, 2), (x, 1), (y, 2), (y, 1), (z, 1), (z, 2),$$

$$R_2^\alpha(S_2 \times \{1, 2\})\colon (x, 1), (x, 2), (y, 1), (y, 2), (z, 2), (z, 1),$$

$$\forall i \in N \backslash \{1, 2\}\colon \tilde{R}_i^\alpha(S_2 \times \{1, 2\}) = \tilde{R}_1^\alpha(S_2 \times \{1, 2\}),$$

where $\tilde{R}_j^\alpha(S_2 \times \{1, 2\}) = \tilde{R}_j^\alpha \cap [(S_2 \times \{1, 2\}) \times (S_2 \times \{1, 2\})]$ for all $j \in N$, and that

$$\forall (v, j) \in (X \times N) \backslash (S_2 \times \{1, 2\})\colon \begin{cases} ((z, 2), (v, j)) \in P(\tilde{R}_1^\alpha), \\ ((z, 1), (v, j)) \in P(\tilde{R}_2^\alpha), \\ \forall i \in N \backslash \{1, 2\}\colon ((v, j), (x, 2)) \in P(\tilde{R}_i^\alpha), \end{cases}$$

$$\forall i \in N, \forall (v^1, j^1), (v^2, j^2) \in (X \times N) \backslash (S_2 \times \{1, 2\})\colon$$
$$((v^1, j^1), (v^2, j^2)) \in I(\tilde{R}_i^\alpha).$$

It is easy to verify that this profile satisfies the axiom of identity. Note that $E_q^\alpha(S_2) = \{z\}$ and $P(R_N^\alpha) \cap (S_2 \times S_2) = \{(x, y), (y, z), (x, z)\}$, so that we have $C^\alpha(S_1) = \{x\}$ by virtue of CBEP. Consider now $C^\alpha(S_2)$. If x or y belongs to $C^\alpha(S_2)$, then $z \in C^\alpha(S_2)$ by virtue of FI(a). If $z \in C^\alpha(S_2)$, then x as well as y belongs to $C^\alpha(S_2)$ thanks to FI(b). $C^\alpha(S_2)$ being non-empty, we should then conclude that $C^\alpha(S_2) = S_2$. Then we have $S_1 \subset S_2$, $S_1 \cap C^\alpha(S_2) = \{x, y\} \not\subset C^\alpha(S_1) = \{x\}$, which implies that a GCCR satisfying UID, FI and CBEP cannot possibly satisfy CA. ∎

3.5. Several remarks seem to be in order here.

Firstly let us point out that there exists a concrete example of fairness-extending as well as fairness-including GCCR which is due essentially to Goldman and Sussangkarn [8]. For any profile $\alpha \in \mathscr{A}$ and any $S \in \mathscr{S}$, let a binary relation R^α on X be defined by $R^\alpha = P(R_N^\alpha) \cup R_E^\alpha$ and let $R^\alpha(S)$ stand for the restriction of R^α on $S: R^\alpha(S) = R^\alpha \cap (S \times S)$. We then define a choice set $C^\alpha_{\mathrm{GS}}(S)$ for S by

$$C^\alpha_{\mathrm{GS}}(S) = \{x \in S \mid \forall y \in S: (x, y) \in T(R^\alpha(S)) \vee (y, x) \notin T(R^\alpha(S))\}, \quad (13)$$

where $T(R^\alpha(S))$ denotes the transitive closure of $R^\alpha(S)$. Associating the well-defined choice function C^α_{GS} on $\mathscr{S}$ thereby constructed with the profile α we have started from, we have a complete description of a GCCR Ψ_{GS}: $\Psi_{\mathrm{GS}}(\alpha) = C^\alpha_{\mathrm{GS}}$ for all $\alpha \in \mathscr{A}$. It is easy, if tedious, to show that Ψ_{GS} satisfies the condition FE as well as the condition FI for all $\alpha \in \mathscr{A}$. Thanks to Theorem 2 we can assert without further ado that Ψ_{GS} cannot possibly satisfy the superset axiom of choice-consistency. We may also prove that Ψ_{GS} fails to satisfy Chernoff's axiom of choice-consistency as well, but it does satisfy the stability axiom, although the latter property may presumably be too weak to celebrate Ψ_{GS} for its success in this arena. The power of Theorem 2 is such that it asserts by one stroke that we cannot possibly improve the performance of Ψ_{GS} unless we renounce the wide applicability of our GCCR or the nice choice-consistency property thereof.

Secondly we may assert the following simple corollaries of Theorems 1, 2, and 3. In view of the strong intuitive appeal of path-independence argument, these corollaries may better crystallize the logical difficulty identified by Theorems 2 and 3.

COROLLARY 1. *Suppose that there exist at least three social states. Then there exists no GCCR which satisfies UID (Unrestricted Domain under the Axiom of Identity), FE (Fairness Extension) and WPI(β) (Weak Path-Independence β).*

COROLLARY 2. *Suppose that there exist at least three social states. Then there exists no GCCR which satisfies UID (Unrestricted Domain under the Axiom of Identity), FI (Fairness Inclusion), CBEP (Conditional Binary Exclusion Pareto) and WPI(α) (Weak Path-Independence α).*

Thirdly we should note that Theorem 2 as well as Theorem 3 do not invoke any interprofile independence condition, which has often been nominated as *the* culprit of Arrovian impossibility theorems. Indeed, only a single profile is made effective use of in proving Theorems 2 and 3, respectively, so that the *profiles richness condition* UID is in fact much stronger than is needed.[8] Instead of requiring UID we may do throughout with the following *states richness condition* suggested by Pollak [13] and Roberts

---

[8] In this respect, our Theorem 2 and Theorem 3 are similar in nature to Sen's [16, Chap. 6*] *impossibility of a Paretian liberal.*

[15], which is the single-profile analogue of the multiple-profile requirement UID. Let $\alpha \in \mathscr{A}$ denote the given fixed profile.

STATES RICHNESS CONDITION (SRC). *Let* $\beta(S_0)$ *denote any logically possible sub-profile over the hypothetical triple set* S_0. *Then there exists a one-to-one correspondence* γ_β *from* S_0 *into* X *such that*

$$((\gamma_\beta(x), i), (\gamma_\beta(y), j)) \in \tilde{R}^\alpha_k \leftrightarrow ((x, i), (y, j)) \in \tilde{R}^\beta_k$$

for all $x, y \in S_0$ *and all* $i, j, k \in N$.[9]

It should be clear that the condition SRC, which essentially requires that the set of states X is rich enough, may replace the condition UID, which is the requirement to the effect that the set of profiles $\mathscr{A}$ is rich enough, to generate single-profile analogue of Theorems 2, 3 and Corollaries 1, and 2.

4. CONCLUDING REMARKS

It is hoped that our results reported in this paper, which are largely negative, will help clarify the nature and potentiality of the fairness-as-no-envy approach in the theory of fairness and justice. In concluding this paper, a few remarks are due.

(a) According to Varian [20, p. 65], "[s]ocial decision theory asks for too much out of the [preference aggregation] process in that it asks for an entire *ordering* of the various social states.... The original question asked only for a good allocation; there was no requirement to rank all allocations. The fairness criterion in fact limits itself to answering the original question. It is limited in that it gives no indication of the merits of two nonfair allocations, but by restricting itself in this way it allows for a reasonable solution to the original problem." This contrast between "social decision theory" and "fairness criterion" is no doubt a useful one, but it seems to us that the two approaches may well be subsumed in a more general choice-functional collective choice framework. In doing so, we may enrich our understanding of one theory in the light of the implications of the other theory on the common ground and *vice versa*. This is precisely the kind of exercise we tried to perform in this paper.

(b) It is often suggested that the prime virtue of the theory of fairness is that it requires no such things as *externally imposed* interpersonal welfare comparisons, hypothetical welfare functions, or fictitious original position. Notice that our analysis of the concept of fairness in the framework of social choice theory fully retains this alleged prime virtue of the theory of fairness. One may even claim, following Alchian [1], that what is involved in our GCCR framework is not an interpersonal welfare comparison but an intrapersonal, intersituational comparison.

---

[9] More explicitly, $\beta(S_0) = (\tilde{R}^\beta_1(S_0 \times N), \tilde{R}^\beta_2(S_0 \times N), \dots, \tilde{R}^\beta_n(S_0 \times N))$.

REFERENCES

1. A. A. ALCHIAN, The meaning of utility measurement, *Amer. Econ. Rev.* **43** (1953), 26–50.
2. K. J. ARROW, Rational choice functions and orderings, *Economica* **26** (1959), 121–127.
3. K. J. ARROW, "Social Choice and Individual Values," 2nd ed., Wiley, New York, 1963.
4. D. H. BLAIR, G. BORDES, J. S. KELLY, AND K. SUZUMURA, Impossibility theorems without collective rationality, *J. Econ. Theory* **13** (1976), 361–379.
5. T. E. DANIEL, A revised concept of distributional equity, *J. Econ. Theory* **11** (1975), 94 109.
6. J. A. FEREJOHN AND D. M. GRETHER, Weak path independence, *J. Econ. Theory* **14** (1977), 19–31.
7. D. K. FOLEY, Resource allocation and the public sector, *Yale Econ. Essays* **7** (1967), 45–98.
8. S. M. GOLDMAN AND C. SUSSANGKARN, On the concept of fairness, *J. Econ. Theory* **19** (1978), 210–216.
9. E. A. PAZNER, Pitfalls in the theory of fairness, *J. Econ. Theory* **14** (1977), 458–466.
10. E. A. PAZNER AND D. SCHMEIDLER, A difficulty in the concept of fairness, *Rev. Econ. Studies* **41** (1974), 441–443.
11. E. A. PAZNER AND D. SCHMEIDLER, Egalitarian equivalent allocations: A new concept of economic equity, *Quart. J. Econ.* **92** (1978), 671–687.
12. C. R. PLOTT, Path independence, rationality, and social choice, *Econometrica* **41** (1973), 1075–1091.
13. R. A. POLLAK, Bergson-Samuelson social welfare functions and the theory of social choice, *Quart. J. Econ.* **93** (1979), 73–90.
14. J. RAWLS, "A Theory of Justice," Harvard Univ. Press, Cambridge, Mass., 1971.
15. K. W. S. ROBERTS, Social choice theory: The single-profile and multiprofile approaches, *Rev. Econ. Studies* **47** (1980), 441–450.
16. A. K. SEN, "Collective Choice and Social Welfare," Holden–Day, San Francisco, 1970
17. A. K. SEN, Social choice theory, *in* "Handbook of Mathematical Economics" (K. J. Arrow and M. Intriligator, Eds.), North-Holland, Amsterdam, in press.
18. P. SUPPES, Some formal models of grading principles, *Synthese* **6** (1966), 284–306.
19. K. SUZUMURA, "Equity, Utility and Majority: Extended Sympathy Approach," Department of Economics, Stanford University, July 1980.
20. H. R. VARIAN, Equity, envy and efficiency, *J. Econ. Theory* **9** (1974), 63–91.
21. H. R. VARIAN, Distributive justice, welfare economics, and the theory of fairness, *Philos. Pub. Affairs* **4** (1975), 223–247.
22. H. R. VARIAN, Two problems in the theory of fairness, *J. Pub. Econ.* **5** (1976), 249–260.

JOURNAL OF ECONOMIC THEORY 25, 380–396 (1981)

Self-Fulfilling Prophecies

COSTAS AZARIADIS*

Department of Economics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received April 14, 1980; revised October 31, 1980

After a quarter century of unbroken development in the theory of allocation under uncertainty, it has become an obvious fact that randomness in endowments, preferences or technology will generally work its way to the prices and allocations which prevail in equilibrium. Is it true, as intuition may suggest in haste, that random prices necessarily reflect some intrinsic uncertainty in the structure of the economy, or can they arise, as some recent literature [3, 12, 14, 15] indicates, merely from extraneous, self-perpetuating beliefs that prices are stochastic?

The question is of interest for it raises the possibility that business cycles are set in motion by arbitrary shifts in any factor, however purely subjective, agents happen to deem relevant to economic activity: animal spirits, consumer sentiment or the prophecies of the Sibyl at Cumae may spark fluctuations in which prices change simply because they are expected to and price signals convey no structural information.

The evidence on the influence of subjective factors is ample and dates back several centuries[1]; the Dutch "tulip mania," the South Sea bubble in England, and the collapse of the Mississippi Company in France are three well-documented cases of speculative price movements which historians consider unwarranted by "objective" conditions.

What follows is a demonstration that a kindred type of paradoxical behavior, which we name *extraneous uncertainty*, is both possible and "frequent" among rational expectations equilibria in an aggregative model of overlapping generations. In particular, if we constrain (the probability distribution of) the price level to clear markets, reproduce beliefs and, in

---

* My interest in this delphic topic stems from conservations with Karl Shell who doesn't necessarily agree with the outcome. Financial support from the National Science Foundation (under Grants SOC 78-00549 and SES 80-06236) and from the Center for the Study of Organizational Innovation at Pennsylvania is acknowledged with thanks, and so is the benefit of Alan Blinder's classical learning. Olivier Blanchard, David Cass, Jerry Green, Menahem Yaari and an eponymous referee gave me helpful comments.

[1] See, for example, [4, Chap. 10]. Although this influence has long been recognized in the oral tradition of economics [6, p. 204], it has not to my knowledge received much formal attention until Karl Shell's recent example in [12].

0022-0531/81/060380–17$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

addition, to follow a simple two-state Markov process, then a significant fraction of the resulting equilibria suffer from extraneous uncertainty. We give some examples of self-fulfilling prophecies that correspond to a permanent "recession" and permanent "boom" in an economy entirely free from price rigidities.

This phenomenon appears to be robust to changes in preferences and in the store of value: it will exist, for instance, under a gross substitutes assumption or a replacement of fiat money with unconditional claims on productive assets. The effect of extraneous uncertainty disappears, however, when the horizon is truncated; in the stationary state this may occur as well if enough contingent claims markets are appended to the economy.

I. PERFECT FORESIGHT

The most convenient way to introduce the basic structure and fix notation is to review the dynamics of the perfect foresight case. We begin with the simplest possible overlapping-generations model[2]: Time extends from zero to infinity; at the beginning of each time period $t = 1, 2, \ldots$, a fixed-size generation of identical individuals is born which lives for two periods, youth and old age. Consumption occurs only in old age, production takes place solely in youth. The very first generation, born at $t = 0$, is "old": each member of it is endowed with one unit of fiat money. Each member of the generation currently young may use a constant-returns-to-scale technology to transform n units of his own leisure into $y \leqslant n$ units of a perishable good for which he has no immediate use. The young, endowed with one unit of divisible leisure each, are obviously motivated to trade goods for the fiat store of value which is held exclusively by the old.

For each generation, preferences over current work and future consumption are additive, viz.,

$$v_t(c_{t+1}, n_t) = u(c_{t+1}) - g(n_t), \tag{1}$$

where $u(\cdot)$ and $-g(\cdot)$ are smooth monotone, concave functions. Let p_t, y_t be the price level and goods supply per young person in period t. The pair of sequences $\{p_t^*\}_0^\infty$ $\{y_t^*\}_0^\infty$ is an *equilibrium*, if:

(a) For each $t \geqslant 1$, y_t^* is indeed equal to the amount of goods supplied by each young person, given the price-level sequence $\{p_t^*\}_0^\infty$. In other words, y_t^* maximizes $\{u(c_{t+1}) - g(y_t)\}$ s.t. $c_{t+1} = p_t^* y_t / p_{t+1}^*$ and $0 \leqslant y_t \leqslant 1$;

(b) the initial condition y_0^* is feasible, i.e., lies in the interval $[0, 1]$; and

---

[2] More general treatments appear in [10, 11, 5]. Readers familiar with this literature may skim Section I to pick up the notation.

(c) the demand for nominal money balances per head equals the corresponding supply, i.e., $p_t^* y_t^* = 1$ for all t.

We begin our examination of perfect foresight dynamics quite informally. Let us define $s(p_t/p_{t+1})$ as the solution to the maximum problem in (a) above, i.e., as the apount of goods supplied by each young person in period t if current price is p_t and future price is p_{t+1}. Clearly $s(\cdot)$ is a single-valued function about which we cannot say much unless we are willing to restrict preferences somewhat. For instance, $s(\cdot)$ is monotone increasing (decreasing) if current leisure and future consumption are gross substitutes (gross complements) at all values of the price ratio p_t/p_{t+1}.

The market clears if the price sequence $\{p_t\}_0^\infty$ satisfies

$$s(p_t/p_{t+1}) = 1/p_t. \tag{1a}$$

This first-order difference equation, together with the price level p_0 in period zero, describes the evolution of equilibrium prices over time.

Equation (1a) clearly reveals that the production economy we are studying in this paper is very close to the simple pure exchange economy frequently studied in the overlapping-generations literature. Suppose, for instance, that each member of generations $t = 1, 2, \ldots,$ were endowed not with leisure but with positive amounts (e_1, e_2) of a perishable consumption good in youth and old age. Given this endowment vector, let $s(p_t/p_{t+1})$ be now what the young *save* at the price ratio p_t/p_{t+1} and assume that $s(p_t/p_{t+1}) > 0$, for all values of p_t/p_{t+1}. The dynamics of this exchange model is again given by Eq. (1a).

Returning to our production economy, we define a *stationary* equilibrium price sequence as any non-negative constant sequence $\{p^*\}_0^\infty$ satisfying Eq. (1a). Clearly, $p_t = \infty$ for all t is one such sequence and it supports an autarkic equilibrium with zero output and worthless money.

There is exactly *one* other stationary equilibrium price sequence with valuable money, i.e., such that $p^* < \infty$. This one is defined from

$$s(1) = 1/p^*; \tag{1b}$$

it is unique because $s(\cdot)$ is single-valued, i.e., only one value of goods supply corresponds to the maximal element of the budget set $\{y_t, c_{t+1} \mid 0 \leqslant c_{t+1} \leqslant y_t\}$.

The stability of this particular stationary state depends very much on preferences. Suppose, for instance, that consumption and leisure are gross substitutes and let $p_t > p^*$ for some t. Then $1/p_t < 1/p^*$, and equilibrium requires that $s(p_t/p_{t+1}) < s(1)$. Since $s(\cdot)$ is increasing, we have $p_t/p_{t+1} < 1$ or $p_{t+1} > p_t$, thus moving further away from p^*. Gross substitutability in our simple perfect-forsight economy implies that the stationary state with finite price level is unstable.

More formally, if we define two functions

$$G(n) \equiv ng'(n), \qquad U(c) \equiv cu'(c) \tag{2}$$

such that $G(n) \to 0$ as $n \to 0$, assume that $G(n) \to \infty$ as $n \to 1$, and solve the maximum problem in part (a) of the previous definition, we find that every equilibrium sequence satisfies three requirements: the initial condition $y_0 = y_0^*$; feasibility, i.e., $y_t \in [0, 1]$; and a "law of motion," viz.,

$$\begin{aligned} y_{t+1} &= 0, &\quad \text{if} \quad y_t &= 0, \\ U(y_{t+1}) &= G(y_t), &\quad \text{if} \quad y_t &> 0. \end{aligned} \tag{3}$$

We note that, in equilibrium, y_t equals the commodity price of money in period t; Eq. (3) is merely Eq. (1a) with preferences made explicit.

For stationary equilibria, Eq. (3) becomes

$$U(y^{**}) = G(y^{**}). \tag{3$'$}$$

It is well known [2, 13] that these three requirements are not sufficient to determine the price level, p_t (equivalently, the commodity price of money) unless we know the initial price $y_0 = 1/p_0$. To each value of y_0 there corresponds typically a different equilibrium sequence $\{y_t^*\}_0^\infty$. To see this we graph Eq. (3) in Fig. 1 and confirm readily that there exist two steady states: one corresponding to zero price of money (at the origin), another to a positive price (at y^{**}). In panel III, for example, the stationary state S with positive price of money is stable and may be reached in infinite time from any $y_0 \in (0, 1)$. Hence, there is an infinity of equilibrium price sequences.

Initial conditions are, of course, arbitrary; our model provides the economy with no mechanism to choose among them, leaving the equilibrium price sequence indeterminate. Even if we somehow confined our attention to equilibria which are not Pareto-dominated by other equilibria,[3] it is not obvious that we would be left with a unique solution since many of these multiple equilibria turn out to be Pareto-noncomparable. An illustration of this dilemma is in panel IV of Fig. 1, which corresponds to preferences $g(n) = n/k^2$, $u(c) = -c^{-1}$ for some constant $k \in (0, 1)$. Then $G(n) = n/k^2$, $U(c) = 1/c$, and (3) yields

$$y_{t+1}\, y_t = k^2. \tag{4}$$

---

[3] The suggestion is discussed in [13, 1].

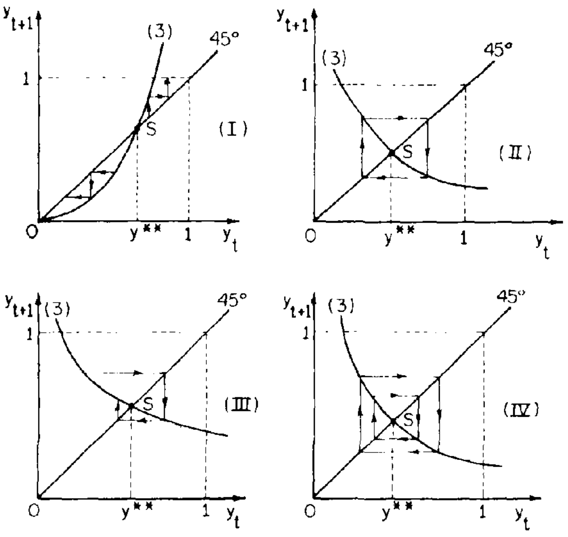


FIGURE 1

The non-trivial stationary solution is $y^{**} = k$ but there are others for which the economy cycles, e.g.,

$$y_t^* = a, \qquad y_{t+1}^* = k^2/a, \tag{5}$$

for any constant $a \in [k^2, 1]$. The non-trivial stationary solution yields utility $-2/k$ for each generation whereas the cycling solution yields $(-2a/k^2, -2/a)$ to successive generations. For $a \neq k$, one sees easily that

$$\min(-2a/k^2, -2/a) < -2/k < \max(-2a/k^2, -2a) \tag{6}$$

and, hence, (5) cannot be compared with $y^{**} = k$.

To sum up: The non-uniqueness of the equilibrium price level in the overlapping generations model is a phenomenon which owes much to the infinite number of decision makers and dated commodities, and very little to properties like lack of gross substitutability which are crucial in a static general equilibrium context.[4]

What is less well understood is that extraneous uncertainty considerably enlarges the set of equilibrium prices that arise under perfect foresight, by adding in certain cases an infinity of self-replicating equilibria. Some of these are examined next.

---

[4] This point, also made by Shell *et al.* [13] and Calvo [2, Sect. II], is evident in panel I, Fig. 1. The price level sequence $\{p_t^*\}_0^\infty$ is generally indeterminate there *despite* the fact that the function U is increasing and, hence, current leisure and future consumption are gross substitutes.

II. Replicating Equilibria

We can now introduce extraneous uncertainty at little cost in additional notation. Denote by Ω_t the information set available to agents in period t; this may include any historical element in the economy such as prices and quantities. Let ω_t be a typical element of Ω_t. By analogy with the preceding section, *a pair of sequences of random variables* $\{\tilde{p}_t\}_0^\infty$ $\{\tilde{y}_t\}_0^\infty$ *is an equilibrium if*:

(a) For each ω_t and $t \geqslant 1$, $\tilde{y}_t = y(\omega_t)$ solves

$$\max\{E[u(c_{t+1}) \mid \Omega_t] - g(y_t)\} \qquad \text{s.t. } c_{t+1} = \tilde{p}_t y_t / \tilde{p}_{t+1};$$
$$0 \leqslant \tilde{y}_t \leqslant 1;$$

(b) $\tilde{y}_0$ is a number in $[0, 1]$; and

(c) $\tilde{p}_t \tilde{y}_t = 1$ for all ω_t.

At the individual level, the probability distribution of $\tilde{y}_t$ is derived from that of $\tilde{p}_t$ and both are conditional on Ω_t. In the aggregate, the "law of motion" for the economy becomes

$$E\{U(y_{t+1}) \mid \Omega_t\} = G(y_t), \tag{7}$$

the solution being again an appropriate conditional probability distribution which confirms price expectations. This law, however, does not constrain sufficiently the solution for it says nothing whatever about the higher moments of the random variable $U(y_{t+1})$. Suppose, for example, that ε_t is *any* independent, identically distributed random variable with mean one, belonging to the information set Ω_t; then the stochastic difference equation

$$U(y_{t+1}) = G(y_t)/\varepsilon_{t+1} \tag{8}$$

"solves" Eq. (7) and corresponds to a rational expectations equilibrium if, given y_t, ε_{t+1} is defined so that $y_{t+1} \in [0, 1]$ with probability 1 for all t.

To avoid problems with feasibility and, at the same time, restrict somewhat the equilibrium price set, we limit ourselves to solutions with the Markov property, i.e., ones for which

$$\Omega_t = y_t. \tag{9}$$

In particular, suppose that, for all t, the price of money may attain at most two values, $y_t = \{y_1, y_2\}$, with the following stationary transition probability matrix:

$$\begin{array}{c|cc} y_t \diagdown y_{t+1} & y_1 & y_2 \\ \hline y_1 & q_1 & 1-q_1 \\ y_2 & 1-q_2 & q_2 \end{array} = T. \tag{10}$$

For those who like to think of price randomness as a "structural" phenomenon, this matrix simply reflects the stochastic properties of any variable deemed by public opinion to have a bearing on economic activity. Since there is myriad of such candidate variables, it makes some sense to treat the probabilities (q_1, q_2) as parameters in this economy—just as we do with prices.

A self-fulfilling equilibrium is now a set of four numbers (q_1, q_2, y_1, y_2) *all lying in the interval* $[0, 1]$ *and satisfying Eq.* (7), viz.

$$q_1 U(y_1) + (1 - q_1) U(y_2) = G(y_1), \tag{11a}$$

$$(1 - q_2) U(y_1) + q_2 U(y_2) = G(y_2). \tag{11b}$$

Since these are two equations in *four* unknowns,[5] we should generally expect multiple equilibria. Some of them we already know: if the economy has a stationary state y^{**} under perfect foresight, that is, a solution to $G(y) = U(y)$, then every quadruple $(q_1, q_2, y^{**}, y^{**})$ such that $q_1 \in [0, 1]$ and $q_2 \in [0, 1]$ obviously solves (11a)–(11b) and is an equilibrium. Extraneous uncertainty, of course, involves *different* prices of money in the two states.

There are at least two special assumptions that will exclude extraneous uncertainty in this self-replicating example. The more obvious one is that consumption and leisure are gross substitutes, that is, U is an increasing function of y. Then, under perfect foresight, we recall that no equilibrium price sequence goes to the non-trivial stationary state unless it started there (see also Fig. 1, panel I). A similar problem appears under uncertainty: goods supply is increasing in the price level, a fact which is not consistent with market clearing, i.e., with

$$p_1 y_1 = p_2 y_2 = 1. \tag{12}$$

Formally, solve (11a)–(11b) for (q_1, q_2) to obtain

$$q_1 = [U(y_2) - G(y_1)]/[U(y_2) - U(y_1)], \tag{13a}$$

---

[5] With N states if nature, we would have N equations in N^2 unknowns.

$$q_2 = [G(y_2) - U(y_1)]/[U(y_2) - U(y_1)]. \tag{13b}$$

Suppose, without any loss of generality, that $y_2 > y_1$, so that $U(y_2) > U(y_1)$. Then $q_1 < 1$ $(\Rightarrow)$ $G(y_1) > U(y_1)$ $(\Rightarrow)$ $g'(y_1) > u'(y_1)$ $(\Rightarrow)$ $y_1 > y^{**}$. Similarly $q_2 < 1$ $(\Rightarrow)$ $G(y_2) < U(y_2)$ $(\Rightarrow)$ $g'(y_2) < u'(y_2)$ $(\Rightarrow)$ $y_2 < y^{**}$. Hence $y_1 > y_2$, which contradicts the maintained assumption $y_2 > y_1$. The only solution is $y_1 = y_2 = y^{**}$.

The same situation arises under gross complementarity if $q_1 + q_2 \geqslant 1$. Then $y_2 > y_1$ implies $U(y_2) < U(y_1)$; from (13a) and (13b), however, $G(y_1) \geqslant G(y_2)$ if $q_1 \geqslant 1 - q_2$ and, hence, $y_1 \geqslant y_2$, which contradicts a maintained assumption, since G is by definition an increasing function. A symmetric contradiction obtains if the maintained assumption is $y_1 > y_2$.

For the remainder of this section we suppose that consumption and leisure are gross complements. If $y_2 > y_1$ then $U(y_2) < U(y_1)$, and q_1 is a probability if, and only if

$$U(y_2) \leqslant G(y_1) \leqslant U(y_1). \tag{13c}$$

Also, $q_2 \in [0, 1]$ if and only if

$$U(y_2) \leqslant G(y_2) \leqslant U(y_1). \tag{13d}$$

Note, however, that $G(y_1) \geqslant U(y_2)$ $(\Rightarrow)$ $G(y_2) \geqslant U(y_2)$ and $U(y_1) \geqslant G(y_2)$ $(\Rightarrow)$ $U(y_1) \geqslant G(y_1)$. Hence, with $y_2 > y_1$, q_1 and q_2 are probabilities if, and only if

$$G(y_1) \geqslant U(y_2); \qquad U(y_1) \geqslant G(y_2). \tag{14a, b}$$

Then one easily verifies that $q_1 + q_2 < 1$. Similarly, with $y_2 < y_1$, the q's are probabilities with sum no more than one if, and only if,

$$G(y_1) \leqslant U(y_2); \qquad U(y_1) \leqslant G(y_2) \tag{14c, d}$$

The equations $G(y_1) = U(y_2)$ and $U(y_1) = G(y_2)$ are clearly symmetric in the plane around the straight line $y_1 = y_2$. In Fig. 2, then, inequalities (14a, b) are satisfied by all points lying simultaneously not below the line $U(y_2) = G(y_1)$, not above the line $U(y_1) = G(y_2)$ *and* above the 45° line; similarly, inequalities (14c, d) are satisfied by all points not above the line $U(y_2) = G(y_1)$, not below $U(y_1) = G(y_2)$ *and* below the 45° line. The set of points that satisfies all *four* inequalities (14a)–(14d) is shaded in Fig. 2. No equilibrium with extraneous uncertainty exists in panel I but there is an infinity of them in panel II. What is more important, the shaded area has significant size (i.e., is not a set of measure zero) relative to the set of all feasible allocations when q_1 and q_2 are both positive. Otherwise, some of the relations in (14) become strict equalities, the shaded area collapses to one of

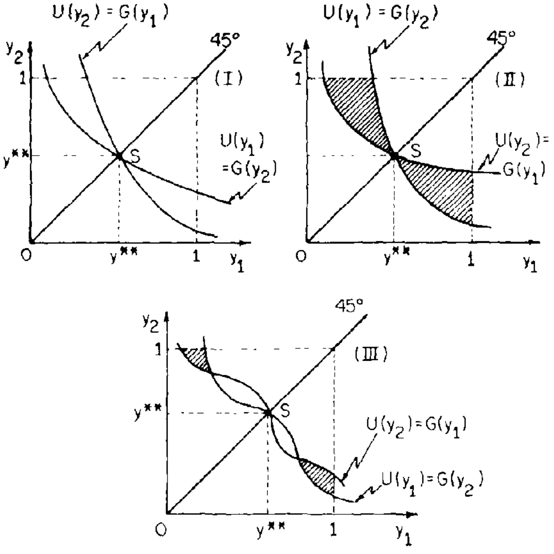


FIGURE 2

its boundaries, and the set of equilibria with extraneous uncertainty has measure zero.

A sufficient condition for panel II to obtain is obviously that the line $U(y_1) = G(y_2)$ be steeper than the line $U(y_2) = G(y_1)$ at point S, or

$$G'(y^{**})/U'(y^{**}) + 1 > 0, \tag{15}$$

i.e., that the "law of motion" in Eq. (3) yield a locally stable nontrivial stationary equilibrium in the perfect foresight case. Panel III in Fig. 2 shows that this condition is not necessary; it is violated at S, and yet extraneous uncertainty appears in the shaded areas near the NW and SE corners.

In summary, if $q_1 + q_2 < 1$, a set of sufficient (but not necessary) conditions that guarantee the existence of replicating equilibria with extraneous uncertainty under the stochastic structure postulated in (10) are: gross complementarity between consumption and leisure, and *local stability* of the non-trivial stationary state. Furthermore, if the transition probability matrix is ergodic, then the set of these equilibria has the same dimension as that of all feasible allocations.

III. ARE SELF-FULFILLING PROPHECIES LIKELY?

Readers who, for intuitive reasons of their own, remain unconvinced of the importance of self-fulfilling prophecies may regard the example in the preceding section as an artifact of the imagination that is unlikely to occur in "practice." Yet, under the assumptions made in the previous paragraph, extraneous uncertainty is not only possible but "probable" as well; for most configurations of the exogenous probabilities, q_1 and q_2 such that $q_1 + q_2 \leqslant 1$, there exists one stationary equilibrium and at least *two* other distinct equilibria such that $y_1 \neq y_2$. To see this, suppose $g'(1) = \infty$, $U(c) \to B \leqslant \infty$ as $c \to 0$, and define the function $\hat{T}$ from

$$\hat{T}(q, y) \equiv (1 - q)^{-1} [G(y) - qU(y)] \tag{16}$$

for $q < 1$. Then $\hat{T}$ is increasing in y, and Eqs. (11a)–(11b) reduce to

$$U(y_2) = \hat{T}(q_1, y_1); \qquad U(y_1) = \hat{T}(q_2, y_2). \tag{17a, b}$$

To solve (17a, b) we require some additional information from (16), namely:

$$\hat{T}(q, y) \to \infty \qquad \text{as} \quad y \to 1; \tag{18a}$$

$$\hat{T}(q, y^{**}) = U(y^{**}), \qquad \forall q; \tag{18b}$$

and

$$\hat{T} \geqslant 0 \qquad \text{for} \quad y \in [\hat{y}(q), 1]. \tag{18c}$$

Here the function $\hat{y}(q)$, defined as the solution to $g'(y) = qu'(y)$, is obviously increasing, because of the gross complements assumption, and such that

$$\hat{y}(1) = y^{**}. \tag{19}$$

From the information in (18a, b, c) we may now draw Eqs. (17a, b) in Fig. 3. By the local stability of the stationary equilibriam, the line $U(y_1) = \hat{T}(q_2, y_2)$ is steeper than $U(y_2) = \hat{T}(q_1, y_1)$ at point S. Hence, to each pair (q_1, q_2) of probabilities such that $q_1 \neq q_2$ corresponds the non-trivial stationary solution y^{**} and, in addition, at least *two* distinct equilibria with extraneous uncertainty: one at E_1 in which $y_1 < y^{**} < y_2$, another at E_2 such that $y_2 < y^{**} < y_1$. When $q_1 = q_2$, points E_1 and E_2 become symmetric about the 45° line and we have a single equilibrium other than S.

As (q_1, q_2) takes on valuess in the set $[0, 1] \times [0, 1]$, i.e., in the square area of Fig. 4, the two lines in Fig. 3 change position, their intersections yielding the set of all equilibrium prices of money. The shaded area in Fig. 4 corresponds to (q_1, q_2) values for which the stationary value y^{**} is the only equilibrium with positive price of money. Extraneous uncertainty will appear for (q_1, q_2) in the unshaded region and, outside the set of measure zero $Q = \{q_1, q_2 \mid q_1 \in [0, 1];\ q_2 \in [0, 1];\ q_1 + q_2 < 1;\ q_1 = q_2\}$,[6] it will charac-

---

[6] The set Q in Fig. 4 is merely the straight-line segment $(0F)$ minus its end-point F.

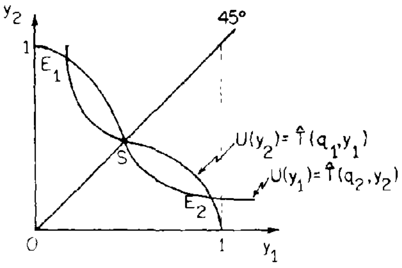


FIGURE 3

terize at least two-thirds of the equilibria in that region. In fact, if there are many intersections in Fig. 3 other than E_1, S, E_2, then nearly *all* the equilibria which correspond to the unshaded area of Fig. 4 will suffer from extraneous uncertainty.

Averaging the two regions of that figure, then, it is clear in this example that self-fulfilling prophecies are at least "one-third" and less than "one-half" of all equilibria.

IV. PERMANENT RECESSIONS

The Markovian structure of the matrix T links the economy's present with its immediate past; in fact, it is fairly straightforward to verify by tedious computation that $q_1 + q_2 < 1$ implies a negative serial correlation in equilibrium output. To ascertain what the economy does over the long haul, one typically considers T^n, the nth power of the matrix T; if T is ergodic, then as $n \to \infty$, T^n tends to some matrix[7]

$$\Pi = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix}, \tag{16}$$

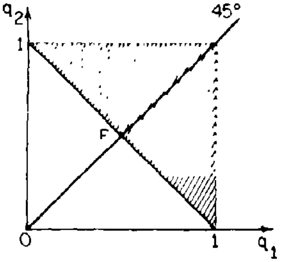


FIGURE 4

---

[7] Cf. [9, p. 197–198].

which obviously satisfies $T \cdot \Pi = \Pi$. The numbers

$$\pi_1 = \frac{1 - q_2}{2 - q_1 - q_2}; \qquad \pi_2 = 1 - \pi_1, \tag{17}$$

are steady-state probabilities that the economy will, in infinite time, find itself in states 1 or 2, irrespective of initial conditions.

The ratio of these probabilities,

$$\pi_1/\pi_2 = (1 - q_2)/(1 - q_1), \tag{18}$$

may attain values arbitrarily close to zero if $q_2 = 1 - 2\varepsilon$, $q_1 = \varepsilon$ and ε is a positive number sufficiently close to zero. Then the economy will either be in the stationary state or "settle" asymptotically on state 2, that is, on a nearly permanent, extreme level of economic activity—recessionary or boomlike.

Since we do not know how a particular equilibrium prevails when several are possible, we cannot be sure that, in some circumstances at least, an extreme equilibrium will obtain. But it is interesting to note that a nearly "permanent" low level of economic activity is consistent with a neoclassical model of equilibrium in which prices are flexible and expectations are rational.

V. ROBUSTNESS

We consider briefly here whether perturbations of the assumptions we employed in the extended example of Sections II–IV will alter the basic message, which is that infinitely many solutions exist with the extraneous uncertainty property.

(a) *Gross Complementarity*

Suppose that consumption and leisure are, instead, gross substitutes. In particular let $u(c) = c$ and $g(n) = n^2$ for $n \in [0, 1]$ so that the law of motion in Eq. (7) becomes

$$2y_t^2 = E(y_{t+1} \mid y_t) \tag{19}$$

with a unique, non-trivial stationary solution, $y^{**} = 1/2$. One verifies immediately that, for any initial $y_0 \in [0, 1/2]$, Eq. (19) is also solved by

$$\begin{aligned} y_{t+1} &= z_{t+1} \qquad \text{w.p.} \quad q_{t+1} \equiv (1 - 4y_t^2)/(1 - 2z_{t+1}) \\ &= 1/2 \qquad \text{w.p.} \quad 1 - q_{t+1}, \end{aligned} \tag{20}$$

where z_{t+1} is any arbitrary number in $[0, 2y_t^2]$. These solutions are infinitely many, because z is picked arbitrarily each period from some interval; and they are all feasible since, for all t, $q_t \in [0, 1]$ and $y_t \in [0, 1]$ w.p. 1. Ergo, there are infinitely many equilibria.

For example, set $z_t = 0$, $\forall t$. Then for any $y_0 \in [0, 1/2]$ and $t \geqslant 1$, we have

$$\begin{aligned} y_t &= 0 \qquad \text{w.p.} \quad 1 - 4y_0^2 \\ &= 1/2 \qquad \text{w.p.} \quad 4y_0^2. \end{aligned} \tag{21}$$

In other words, starting from any point in $(0, 1/2]$, the economy will reach $y^{**} = 1/2$ with positive probability, an event which panel I in Fig. 1 shows to be impossible under perfect foresight.

(b) *The Role of Fiat Money*

Can we blame the inherent worthlessness of fiat money for the great multiplicity oif equilibria which are consistent with self-fulfilling prophecies? If a productive asset like land were the sole store of value,[8] then we know from Calvo [2, Sect. I] that an infinity of equilibrium price sequences are possible under perfect foresight, and we may guess that the situation does not change drastically if one injects prophecies into the economy.

This suspicion is borne out in the following model. Preferences are as in Eq. (1) and there is no money; land is transferred from the old to the young at the beginning of each period in return for consumption goods later in that period. Let N denote hours of work; $Q =$ amount of land; output is $Y = F(Q, N)$ where F is smooth, concave, with constant returns-to-scale. Also, define $y \equiv Y/Q$; $n \equiv N/Q$; $f(z) \equiv F(1, z)$; $p =$ price of land (including rental) in terms of current consumption. Endowments are $\bar{N} = 1$ for leisure and $\bar{Q} = 1$ for land. All variables are per member of the young generation, except $\bar{Q}$ which applies to each member of the old generation.

To purchase one unit of land in period t, the young must work in return $n_t = h(p_t)$ hours, where $h = f^{-1}$ is the inverse production (e.g., labor requirements) function. Each unit of land will entitle the owner to consume p_{t+1} units of output in old age; to entertain the possibility of self-fulfilling prophesis we allow p_{t+1} to be stochastic in principle.

The young demand Q_t^* units of land, which is the solution to

$$\max_{Q_t > 0} E\{U(Q_t p_{t+1}) \mid \Omega_t\} - g[Q_t h(p_t)], \tag{22}$$

with Ω_t being the current information set. At equilibrium $Q_t^* = 1$, $\forall t$, provided $p_t > 0$. Hence, if one defines the increasing function $J(z) \equiv G[h(z)]$

---

[8] Storage of goods over time does not destroy extraneous uncertainty unless the cost of carrying inventories forward is zero. With positive storage cost h per unit, it is possible (see Fig. 2) to pick equilibrium prices of money close enough to the stationary state so that the real rate of return on money always exceeds $-h$.

and goes through the same process as in the beginning of Section II, one finds that every interior rational expectations equilibrium satisfies

$$J(p_t) = E\{U(p_{t+1}) \mid \Omega_t\}. \tag{23}$$

This first-order difference equation is very similar to the one in Eq. (7), permitting again infinitely many stochastic equilibria.

I conclude that it is not fiat money in itself that sustains self-fulfilling prophecies in the economy at hand.

VI. THE ORIGIN OF EXTRANEOUS UNCERTAINTY

To isolate the factors responsible for self-perpetuating prophecies let us examine two changes in the basic model of Section II which are sufficient to do away with the influence of extraneous uncertainty. The first one is an announcement by some authority at $t = 0$ that exchange in period $T > 0$ will be permitted to occur only at some deterministic price p_T (or even at a random price $\tilde{p}_T$, provided that is uncorrelated with p_{T-1}). Then, uncertainty unravels, for Eq. (7) says that, for $t < T$, all equilibrium prices will be deterministic.

By truncating the time horizon, this form of price controls yields finitely many decision makers and steers clear of the large equilibrium price sets which typically arise otherwise.[9] Like many types of price controls, this one assumes the authorities know a lot: given T and the initial condition p_0, the price p_T will support an equilibrium only if (p_0, p_T) are consistent with the law of motion

$$G(p_t^{-1}) = E\{U(p_{t+1}^{-1}) \mid p_t\} \qquad \text{with} \quad p_t^{-1} \equiv 1/p_t, \tag*{)24)}$$

that is, only if authorities have exact knowledge of the dynamic behavior of the economy.

A more important cause of self-fulfilling prophecies seems to be the shortage of claims markets. If the predictions of the Cumaean Sibyl affect allocations in Rome, can the Romans neutralize her interference by building up their financial system and, in particular, by storing value in claims which deliver consumption contingent on what the Sibyl says?

Let q_{ij} be the (stationary) probability that the future state will be j if the current state is i $(i, j = 1, 2)$; π_j = money price of a claim delivering one unit of consumption in state j; z_{ij} = number of claims on state j bought by the old if the past state is i; w_{ij} = number of claims sold by young. Money serves no function anymore, but it is useful to continue thinking of it as a medium of exchange and pretend that each old person has one unit of it.

---

[9] See [8, pp. 111–112; 11].

The old choose z_{ij} by maximizing expected utility of consumption, i.e., solving

$$\begin{gathered}\max q_{i1}u(z_{i1}) + q_{i2}u(z_{i2}) \\ \text{s.t. } z_{i1} \geqslant 0,\ z_{i2} \geqslant 0;\ \pi_1 z_{i1} + \pi_2 z_{i2} \leqslant 1;\ i = 1, 2.\end{gathered} \tag{25a}$$

The young on the other hand, minimize the expected disutility of work needed to acquire any given money revenue, A, from the sale of claims. Thus the w_{ij} solve

$$\begin{gathered}\min q_{i1}\, g(w_{i1}) + q_{i2}\, g(w_{i2}) \\ \text{s.t. } w_{i1} \geqslant 0,\ w_{i2} \geqslant 0;\ \pi_1 w_{i1} + \pi_2 w_{i2} \geqslant A;\ i = 1, 2.\end{gathered} \tag{25b}$$

Furthermore, stationary equilibrium means $A = 1$ with probability one in the money market, and

$$z_{ij} = w_{ij}, \qquad \forall (i, j), \tag{25c}$$

in the market for contingent claims.

For $i = 1, 2$, every interior equilibrium satisfies

$$\frac{\pi_1}{\pi_2} = \frac{q_{i1}u'(z_{i1})}{q_{i2}u'(z_{i2})} = \frac{q_{i1}\, g'(z_{i1})}{q_{i2}\, g'(z_{i2})}, \tag{26a}$$

which implies

$$g'(z_{i1})/u'(z_{i1}) = g'(z_{i2})/u'(z_{i2}). \tag{26b}$$

Hence, the solutions to (25a, b) satisfy

$$z_{i1}^* = z_{i2}^*, \qquad i = 1, 2.$$

Call z_i^* the common value of z_{i1}^*, z_{i2}^*; then, as $A = 1$ for every young person in every state, we have $(\pi_1 + \pi_2) z_1^* = (\pi_1 + \pi_2) z_2^* = 1 \ (\Rightarrow) \ z_1^* = z_2^*$. Consumption (more generally, economic activity) has become entirely independent of the state of nature.

The Sibyl of Cumae now stands thoroughly neutralized; but the vast Roman Empire is blessed with eight more Sibyls, a large number of oracles, religious seers and mystery cults. To render all of them ineffectual requires more claims markets than the Romans can reasonably hope to set up.

Even if we suppose momentarily that all requisite markets were costlessly avilable and open for business, there is no guarantee that they would rob the Sibyl of all power to influence economic events. The present section demonstrates this to occur in a *special case*, that is, when preferences are additive, beliefs are homogeneous and the probability distribution of future prices conditioned on current prices is stationary.

VII. CONCLUSIONS

The central message of this paper is that even perfectly well-behaved economies will typically admit rational expectations equilibria in which the expectations themselves spark fluctuations in the level of business activity. If many individuals naively believe that sunspots[10] or some index of confidence are good predictors of future prices, then they may take actions which tend to bear out their beliefs.

Self-fulfilling prophecies are by their very nature a source of indeterminacy, augmenting appreciably the already very large number of perfect-foresight equilibria which typically emerges in monetary economies with infinitely many agents and commodities. In one fairly extended example (Sections II and III), self-fulfilling prophecies comprise between one third and one-half of all equilibria. Some of these resemble permanent "recessions" or "booms" and all of them are replicating Markov chains, that is, *perpetual* cycles ignited by expectations alone.

Phenomena of this sort will persist if an intrinsically valuable asset like land replaces fiat money as store of value, but *may* unravel if we set up markets for claims contingent on prophecies—whether they *do* unravel is an interesting topic for future research. As a practical matter, however, one ought to recognize that a vast number of financial markets would be needed to neutralize all subjective factors that individuals might consider influential in economic life.

Given some market incompleteness, what can be said in general about the solutions to Eq. (7)? In particular, under what conditions will extraneous uncertainty replicate itself in perpetuity (as it does in the example of Section II), vanish in finite time (as in Eq. (21)) or dissolve asymptotically?

These questions are outside the scope of the paper at hand but appear to be ones that we must face if we wish to characterize intelligently the myriad of equilibria which are consistent with the postulate of rational expectations.

---

[10] Jevons [7, Chaps. VI–VII] attempted with some persistence to generalize and test William Herschel's hypothesis that sunspots influenced economic activity. Jevons' theory, however, is not psychological or expectational; it holds that solar activity affects climatic conditions and, through them, the aggregate production possibility frontier.

REFERENCES

1. O. J. BLANCHARD, Backward and forward solutions for economies with Rational Expectations, *Amer. Econ. Rev.* **69** (Papers and Proceedings, 1979), 114–118.
2. G. A. CALVO, On the indeterminacy of interest rates and wages under Perfect Foresight, *J. Econ. Theory* **19** (1978), 321–337.
3. D CASS AND K. SHELL, In defense of a basic approach, *in* "Models of Monetary Economies" (J. Kareken and N. Wallace, Eds.), pp. 251–260, Federal Reserve Bank of Minneapolis, Minneapolis, Minn, 1980.
4. S. B. CLOUGH AND R. T. RAPP, "European Economic History," 3rd ed., McGraw–Hill, New York, 1975.
5. D. GALE, Pure exchange equilibrium of dynamic economic models, *J. Econ. Theory* **6** (1973), 12–36.
6. J R. HICKS, "Value and Capital," 2nd ed., Oxford Univ. Press, Oxford, 1946.
7. W. S. JEVONS, "Investigations in Currency and Finance," Macmillan & Co., London, 1884.
8. *T. C. KOOPMANS, "Three Essays on the State of Economic Science," McGraw–Hill*, New York, 1957.
9. E. PARZEN, "Stochastic Processes," Holden–Day, San Francisco, 1962.
10. P. A. SAMUELSON, An exact consumption–loan model of interest with or without the social contrivance of money, *J. Pol. Econ.* **66** (1958), 467–482.
11. K. SHELL, Notes on the economics of infinity, *J. Pol. Econ.* **79** (1971), 1002–1011.
12. K. SHELL, Monnaie et allocation intertemporelle, unpublished CNRS memorandum, Paris, 1977.
13. K. SHELL, M. SIDRAUSKI, AND J. E. STIGLITZ, Capital gains, income and savings, *Rev. Econ. Studies* **36** (1969), 15–26.
14. R. J. SHILLER, Rational Expectations and the dynamic structure of macroeconomic models, *J. Monetary Econ.* **4** (1978), 1–44.
15. J. B. TAYLOR, Conditions for unique solutions in stochastic macroeconomic models, *Econometrica* **45** (1977), 1377–1385.

JOURNAL OF ECONOMIC THEORY 25, 397–420 (1981)

Nine Kinds of Quasiconcavity and Concavity*

W. E. DIEWERT

Department of Economics, University of British Columbia, Vancouver, Canada

M. AVRIEL

Faculty of Industrial Engineering, Technion, Haifa, Israel

AND

I. ZANG

Faculty of Management, Tel Aviv University, Tel Aviv, Israel

Received May 14, 1980

1. INTRODUCTION

There is no need to emphasize the importence of quasiconcave (and concave) functions in economics and operations research. However, in view of the many excellent papers and books characterizing quasiconcave functions (e.g., [11; 1; 29; 24, pp. 131–150]), it is perhaps necessary to justify yet another paper on the subject.

It is well known (cf. [11, pp. 87–88] or [30, 27]) that a twice continuously differentiable function f of N variables $x \equiv (x_1, x_2, \ldots, x_N)'$ defined over an open convex set S is concave if and only if the function's Hessian matrix of second order partial derivatives, $\nabla^2 f(x)$, is negative semidefinite for every $x \in S$. This is an extremely useful characterization of a concave function in practice. However, a similar simple characterization for twice differentiable quasiconcave functions had not been reported in the literature. In Section 2 below, we provide relatively simple necessary and sufficient conditions for a twice continuously differentiable function to be quasiconcave.

Our criterion for quasiconcavity in the twice continuously differentiable case can be derived from a criterion involving only local properties of the function. It turns out to be possible to characterize all seven varieties of

---

* This research was supported by a Canada Council grant. We are indebted to C. Blackorby and M. Truchon for helpful comments.

0022-0531/81/060397-24$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

concave and quasiconcave functions which were discussed by Ponstein [29][1] (plus two additional types of concave functions) in a similar manner, using only local properties of the function to be characterized. In Sections 2 to 10 below, we report these "new" characterizations, although in some cases, the "new" characterizations turn out to be fairly close to known characterizations.

To summarize, we hope that the present will serve two purposes: (i) to acquaint economists with the rather extensive operations research literature on quasiconcavity; and (ii) to provide alternative characterizations of the various kinds of concavity in general as well as in the once and twice differentiable cases.

Proofs of new theorems are collected in an appendix.

2. QUASICONCAVITY

DEFINITION 1 (Fenchel [11, p. 117]). A function f of N real variables defined over a convex subset S of R^N is quasiconcave iff the upper level sets $L(\alpha) \equiv \{x: f(x) \geqslant \alpha\}$ are convex sets[2] for all real numbers α.

In the remainder of the paper, we assume that S is a convex subset of R^N. If we require *in addition* that S be an open set, we shall say "open S."

An alternative useful characterization of quasiconcavity is provided by the following result.

THEOREM 1 (Fenchel [11, p. 118]). *f is a quasiconcave function defined over S iff*

$$x^1 \in S, \qquad x^2 \in S, \qquad f(x^1) \leqslant f(x^2), \qquad 0 \leqslant \lambda \leqslant 1,$$

implies

$$f(x^1) \leqslant f(\lambda x^1 + (1-\lambda) x^2).$$

Let $\nabla f(x^0) \equiv (\partial f(x^0)/\partial x_1, ..., \partial f(x^0)/\partial x_N)^T$ denote the column vector of first order partial derivatives of f evaluated at the point $x^0 \equiv (x_1^0, ..., x_N^0)^T$. Let $x^T y$ denote the inner product of the vectors x and y. If the first (second) order partial derivatives of f exist and are continuous functions, we say that f is once (twice) continuously differentiable. If the (two-sided) directional derivative

---

[1] Ponstein actually provided characterizations of convex and quasiconvex functions. However, his results are immediately applicable to concave functions since f is a concave (quasiconcave) function if and only if $-f$ is a convex (quasiconvex) function.

[2] A set S is convex iff for every $x, y \in S$ and scalar λ such that $0 \leqslant \lambda \leqslant 1$, we have $\lambda x + (1-\lambda) y \in S$.

$$D_v f(x^0) \equiv \lim_{t\to 0} [f(x^0+tv) - f(x^0)]/t$$

exists for all feasible directions v (i.e., $v^Tv = 1$ and $x^0 + tv \in S$ for some $t > 0$), then f is *directionally* differentiable at $x \in S$. If in addition, $D_v f(x^0) = v^T\nabla f(x^0)$ for all feasible directions v, then f is (Gateaux) *differentiable* at x^0 (cf. Avriel [2, p. 84]).

THEOREM 2 (Arrow and Enthoven [1, p. 780]). *Let f be a differentiable function defined over an open S. Then f is quasiconcave over S iff $x^1 \in S$, $x^2 \in S, f(x^2) \geqslant f(x^1)$, implies $(x^2 - x^1)^T \nabla f(x^1) \geqslant 0$.*

If f is twice continuously differentiable, the existing literature does not provide us with a simple characterization of quasiconcave functions comparable to the characterization given by Theorem 2 above for the once differentiable case. In order to provide a characterization in the twice differentiable case, it proves to be convenient to first develop a criterion for quasiconcavity in general and then utilize this criterion in the twice differentiable case. However, first we require a preliminary result.

THEOREM 3. *Let f be a quasiconcave function defined over S. Then the minimum of f over any compact (i.e., closed and bounded) line segment contained in S is attained on that line segment; i.e., if we let $x^0 \in S$, $v \in R^N$ with $v^Tv = 1$, $\bar{t}v \in S$, then*

$$\begin{aligned} \alpha \equiv \inf_t\{f(x^0+tv): t\in[0,\bar{t}]\} &= \min_t\{f(x^0+tv): t\in[0,\bar{t}]\} \\ &= f(x^0+t^*v), \end{aligned} \tag{1}$$

where $t^* \in [0, \bar{t}]$ and $[0, \bar{t}]$ denotes the closed line segment joining 0 and $\bar{t}$.

If (1) is true, then we say that f has the *line segment minimum property*. Theorem 3 says that a quasiconcave function has this property.

Before we can provide our third characterization of quasiconcavity (Theorems 1 and 2 provide the first two characterizations), we need to introduce a special type of local minimum that is stronger than the concept of a local minimum but weaker than a strict local minimal. Let g be a function of one variable defined over the closed interval $[b, c]$ in the following three definitions, with $b < c$.

DEFINITION 2. g attains a *one-sided semistrict local minimum from above* at $t_0 \in [b, c)$ (the line segment joining b and c excluding c) iff there exists $\varepsilon > 0$ such that $g(t_0) \leqslant g(t_0 + h)$ for all h such that $0 < h \leqslant \varepsilon \leqslant c - t_0$ and $g(t_0) < g(t_0 + \varepsilon)$.

DEFINITION 3. g attains a *one-sided semistrict local minimum from below* at $t_0 \in (b, c|$ (the line segment joining b and c excluding b) iff there exists $\varepsilon > 0$ such that $g(t_0) \leqslant g(t_0 - h)$ for all h such that $0 < h \leqslant \varepsilon \leqslant t_0 - b$ and $g(t_0) < g(t_0 - \varepsilon)$.

Note that a one-sided semistrict local minimum is not a local minimum since the usual definition of a local minimum at t_0 requires that g be locally minimized in *both* directions at t_0.

DEFINITION 4. g attains a *semistrict local minimum* at $t^0 \in (b, c)$ (the line segment joining b and c excluding both end points) iff g attains a one-sided semistrict local minimum from above above and below at t_0.

Thus g attains a semistrict local minimum at t_0 iff it attains a local minimum at t_0 and the function eventually increases as we travel away from t_0 in either direction. Note that a strict local minimum is a semistrict local minimum which in turn is a local minimum.

THEOREM 4. *A function f defined over S is quasiconcave iff f satisfies the line segment minimum property* (1) *over S and the following property*:

$$x^0 \in S, \qquad v^T v = 1, \qquad \bar{t} > 0, \qquad x^0 + \bar{t} v \in S \textit{ implies that}$$
$$g(t) \equiv f(x^0 + tv)$$
$$\textit{does not attain a semistrict local minimum at any } t \in (0, \bar{t}). \tag{2}$$

COROLLARY 4.1. *A continuous from below* (*lower semicontinuous*) *or a continuous function f defined over S is quasiconcave iff f satisfies* (2).

A function f is continuous from below over S iff $\{x: f(x) \leqslant \alpha\}$ is a closed set for each real number α.

The corollary follows from the theorem since a continuous from below or continuous function satisfies the line segment minimum property (1).

COROLLARY 4.2. *A directionally differentiable function f defined over S is quasiconcave iff*

$$\begin{aligned} &x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t} v \in S, x^0 - \bar{t} v \in S, D_v f(x^0) \\ &\quad \equiv \lim_{h \to 0} [f(x^0 + hv) - f(x^0)]/h = 0 \textit{ implies } g(t) \\ &\quad \equiv f(x^0 + tv) \textit{ defined for } x^0 + tv \in S \textit{ does not attain a semistrict} \\ &\qquad \textit{local minimum at } t = 0. \end{aligned} \tag{3}$$

The corollary follows from the theorem since the directional differentiability of f implies that f is also continuous along any line segment contained in S. Also if $g(t) \equiv f(x^0 + tv)$ attains a semistrict local minimum at $t = 0$, then g must attain a local minimum at $t = 0$. Since f is directionally differentiable, we must have $g'(0) = D_v f(x^0) = 0$.

Corollary 4.3. *A differentiable function f defined over an open S is quasiconcave iff*

$$\begin{aligned} &x^0 \in S,\ v^T v = 1,\ v^T \nabla f(x^0) = 0 \textit{ implies } g(t) \equiv f(x^0 + tv) \\ &\qquad \textit{does not attain a semistrict local minimum at } t = 0. \end{aligned} \tag{4}$$

Corollary 4.4. *A twice continuously differentiable function f defined over an open S is quasiconcave iff*

$$\begin{aligned} &x^0 \in S,\ v^T v = 1,\ v^T \nabla f(x^0) = 0 \textit{ implies } (\mathrm{i})\ v^T \nabla^2 f(x^0)\, v < 0 \\ &\qquad \textit{or } (\mathrm{ii})\ v^T \nabla^2 f(x^0)\, v = 0 \textit{ and } g(t) \equiv f(x^0 + tv) \textit{ does not} \\ &\qquad \textit{attain a semistrict local minimum at } t = 0. \end{aligned} \tag{5}$$

Crouzeix [5] has shown that another characterization of quasiconcavity can be obtained if (5)(ii) is replaced by: $v^T \nabla^2 f(x^0)\, v = 0$ and $g(t) \equiv f(x^0 + tv)$ is a quasiconcave function of t for all t such that $x^0 + tv \in S$.

It is possible to rewrite (5) in a format which is more readily checked if we make use of eigenvalues and restricted eigenvalues (see Samuelson [33, pp. 373–377] or Diewert and Woodland [10, pp. 393–394]). Conditions (6) and (7) together are equivalent to (5) above.

$$\begin{aligned} &x^0 \in S,\ \nabla f(x^0) = 0_N,\ \nabla^2 f(x^0)\, v^i = \lambda_i v^i \text{ for } i = 1, 2, \ldots, N \\ &\qquad \text{with } v^{iT} v^j = \{1 \text{ if } i = j,\ 0 \text{ if } i \neq j\} \text{ implies for each } i, \\ &\qquad \textit{either } (\mathrm{i})\ \lambda_i < 0 \textit{ or } (\mathrm{ii})\ \lambda_i = 0 \text{ and } g^i(t) \equiv f(x^0 + tv^i) \\ &\qquad \text{does not attain a semistrict local minimum at } t = 0; \end{aligned} \tag{6}$$

$$\begin{aligned} &x^0 \in S,\ \nabla f(x^0) \neq 0_N,\ \nabla^2 f(x^0)\, v^i + \nabla f(x^0)\, k_i = \lambda_i v^i \text{ for } i = 1, 2, \ldots, N-1 \\ &\qquad \text{with } v^{iT} \nabla f(x^0) = 0 \text{ for } i = 1, 2, \ldots, N-1 \text{ and} \\ &\qquad v^{iT} v^j = \{1 \text{ if } i = j,\ 0 \text{ if } i \neq j\} \text{ implies for each } i, \\ &\qquad i = 1, 2, \ldots, N-1,\ \textit{either } (\mathrm{i})\ \lambda_i < 0 \textit{ or } (\mathrm{ii})\ \lambda_i = 0 \text{ and} \\ &\qquad g^i(t) \equiv f(x^0 + tv^i) \text{ does not attain a semistrict local} \\ &\qquad \text{minimum at } t = 0. \end{aligned} \tag{7}$$

Note that 0_N is a vector of zeroes and the k_i's are scalars. The N λ_i's appearing in (6) are the eigenvalues of $\nabla^2 f(x^0)$; i.e., they are the N roots of the determinantal equation $|\nabla^2 f(x^0) - \lambda I_N| = 0$ and the v^i are corresponding eigenvectors. The $N - 1$ λ_i's appearing in (7) are the eigenvalues of $\nabla^2 f(x^0)$

in the subspace orthogonal to the nonzero gradient vector $\nabla f(x^0)$; i.e., the λ_i's are the $N-1$ roots of the determinantal equation

$$\begin{vmatrix} \nabla^2 f(x^0) - \lambda I_N, & \nabla f(x^0) \\ \nabla^T f(x^0), & 0 \end{vmatrix} = 0 \tag{8}$$

and the v^i appearing in (7) are corresponding restricted eigenvectors.

We turn now to sufficient conditions for quasiconcavity. Obviously, conditions which are sufficient to imply (5) can be obtained by dropping (ii) from (6) and (7). Call the resulting conditions (9) and (10). A condition which is equivalent to (9) is that the Hessian matrix $\nabla^2 f(x^0)$ be negative definite and this latter condition can be translated into well known determinantal conditions (e.g., Bellman [3, p. 74]). Determinantal conditions which are *necessary and sufficient* for (10) (i.e., for $\nabla^2 f(x^0)$ to be negative definite in the subspace orthogonal to $\nabla f(x^0)$) are due to Bellman [3, pp. 77–78], but are much less well known. Finally, the following determinantal conditions are *sufficient* to imply (10) (and hence also (7)):

$x^0 \in S$ implies $(-1)\, B(1, N+1) > 0, (-1)^2\, B(1, 2, N+1) > 0, \ldots,$
$(-1)^N\, B(1, 2, \ldots, N, N+1) > 0$, where

$$B \equiv \begin{bmatrix} \nabla^2 f(x^0), & \nabla f(x^0) \\ \nabla^T f(x^0), & 0 \end{bmatrix}$$

and $B(i_1, i_2, \ldots, i_n)$ denotes the determinant of the submatrix of B consisting of rows and columns $i_1, i_2, \ldots, i_n$. (11)

The restriction $\nabla f(x^0) \neq 0_N$ does not appear in (11) since if $\nabla f(x^0) = 0_N$, then all of the determinants of the submatrices of the bordered Hessian matrix B which include the last column (of zeroes) will be zero. Conditions (11) are the usual (cf. Arrow and Enthoven [1, p. 782]) sufficient conditions for f to be quasiconcave over S. Note that the first inequality in (11) requires that $\partial f(x^0)/\partial x_1 \neq 0$. Note also how much stronger conditions (11) are than the necessary and sufficient conditions (5).

It is obvious that the following condition (compare it with (5)) is a necessary condition for f to be quasiconcave (but it is not in general sufficient):

$$x^0 \in S,\ v^T v = 1,\ v^T \nabla f(x^0) = 0 \text{ implies } v^T \nabla^2 f(x^0)\, v \leqslant 0. \tag{12}$$

It can be shown that condition (12) is equivalent to conditions (13) and (14), where (13) is (6), except that (ii) is dropped and (i) is replaced by the condition $\lambda_i \leqslant 0$ (i.e., the eigenvalues of $\nabla^2 f(x^0)$ are nonpositive), and (14) is (7), except that (ii) is dropped and (i) is replaced by $\lambda_i \leqslant 0$ (i.e., the restricted eigenvalues of $\nabla^2 f(x^0)$ in the subspace orthogonal to $\nabla f(x^0) \neq 0_N$ are nonpositive).

Necessary and sufficient conditions for (13) (i.e., that $\nabla^2 f(x^0)$ be negative semidefinite) are the following determinantal conditions (e.g., Debreu [7]):

$$\begin{aligned} &x^0 \in S, \nabla f(x^0) = 0_N, A \equiv \nabla^2 f(x^0) \text{ implies } (-1) A(i_1) \geqslant 0, \\ &\quad i_1 = 1, 2, \ldots, N; (-1)^2 A(i_1, i_2) > 0, 1 \leqslant i_1 < i_2 \leqslant N; \ldots; \\ &\quad (-1)^N A(i_1, i_2, \ldots, i_N) \geqslant 0, 1 \leqslant i_1 < i_2 < \cdots < i_N \leqslant N, \end{aligned} \tag{15}$$

where $A(i_1, i_2, \ldots, i_n)$ denotes the determinant of the submatrix of A consisting of rows and columns $i_1, i_2, \ldots, i_n$. Note that the determinantal conditions (15) must be checked for every distinct subset of the rows and columns (where the column indices equal the row indices) of $\nabla^2 f(x^0)$. Many economists do not seem to be aware that conditions (15) are necessary and sufficient for the symmetric matrix A to be negative semidefinite, and mistakes are repeatedly made with respect to this point (e.g., Proposition I in [15, p. 597] is false).

Necessary and sufficient conditions for (14) (i.e., that $\nabla^2 f(x^0)$ be negative semidefinite in the subspace orthogonal to $\nabla f(x^0) \neq 0_N$) were derived by Debreu (5) and are equivalent to the following conditions:

$$\begin{aligned} &x^0 \in S, \nabla f(x^0) \neq 0_N \text{ implies } (-1)^2 B(i_1, i_2, N+1) \geqslant 0, \\ &\quad 1 \leqslant i_1 < i_2 \leqslant N; \\ &\quad (-1)^3 B(i_1, i_2, i_3, N+1) \geqslant 0, 1 \leqslant i_1 < i_2 < i_3 \leqslant N; \ldots; \\ &\quad (-1)^N B(i_1, i_2, \ldots, i_N, N+1) \geqslant 0, 1 \leqslant i_1 < i_2 < \cdots < i_N \leqslant N. \end{aligned} \tag{16}$$

The bordered Hessian matrix B and $B(i_1, i_2, \ldots, i_n)$ are defined in (11) above. Note that the conditions $(-1) B(i_1, N+1) = (-1)^2 [\partial f(x^0)/\partial x_1]^2 \geqslant 0$ are automatically satisfied and need not be checked.

Finally, to complete our discussion of determinantal conditions appearing in the literature on quasiconcavity, we note that the following conditions are necessary for conditions (16):

$$\begin{aligned} &x^0 \in S \text{ implies } (-1) B(1, N+1) \geqslant 0, (-1)^2 B(1, 2, N+1) \geqslant 0, \ldots, \\ &\quad (-1)^N B(1, 2, \ldots, N, N+1) \geqslant 0. \end{aligned} \tag{17}$$

Conditions (17) are the usual necessary conditions for f to be a twice continuously differentiable quasiconcave function (cf. [1, p. 781; 27, p. 307; 15, p. 600]).[3]

It is natural to ask what additional conditions can be added to the necessary condition for quasiconcavity (12) in order to obtain sufficiency. Katzner [21, p. 211] shows that (12) is necessary and sufficient for quasiconcavity for the class of functions which have positive first order partial derivatives. The following theorem generalizes Katzner's result.

---

[3] Newman [27, p. 307] and Ginsberg [15, p. 600] call a function f which satisfies conditions (17) a weakly quasiconcave function. This does not appear to be very appropriate terminology, since if $g(x_2, x_3, \ldots, x_N)$ is any twice continuously differentiable function of $N-1$ variables, then $f(x_1, x_2, \ldots, x_N) \equiv g(x_2, \ldots, x_N)$ will be a weakly quasiconcave function of the N variables $x_1, x_2, \ldots, x_N$.

THEOREM 5. *Let f be a twice continuously differentiable function of N variables defined over the open convex set S which satisfies the assumption $\nabla f(x^0) \neq 0_N$ for every $x^0 \in S$. Then f is quasiconcave if and only if condition (12) is satisfied.*

Thus if the twice continuously differentiable function f defined over the open convex set S has a nonzero gradient vector over S, then the determinantal conditions (16) are necessary and sufficient for quasiconcavity.

Crouzeix and Ferland [6] and Ferland [13] have recently obtained additional criteria for quasiconcavity in the twice differentiable case.

3. SEMISTRICT QUASICONCAVITY

DEFINITION 5 (Elkin [14], Ginsberg [15][4]). *f is a semistrictly quasiconcave function defined over S iff*

$$x^1 \in S, \qquad x^2 \in S, \qquad f(x^2) > f(x^1), \qquad 0 < \lambda < 1$$

implies

$$f(\lambda x^1 + (1 - \lambda) x^2) > f(x^1). \tag{18}$$

Semistrict quasiconcavity is important, mainly because of the following theorem.

THEOREM 6 (Martos [25, p. 244], Mangasarian [23, p. 284]). *Let f be a semistrictly quasiconcave function defined over S and suppose f attains a local maximum at $x^0 \in S$. Then f attains a global maximum over S at x^0. Moreover, if f is also continuous from above, then the set of global maximizers is a convex set.*

---

[4] *Mangasarian [23] Ponstein [29], and Thompson and Parke [35] define f to be strictly quasiconcave iff f satisfies (18). However, Karamardian [20] showed that such a strictly quasiconcave function was not necessarily quasiconcave unless the function was also* continuous from above (i.e., unless the upper level sets $L(u) \equiv \{x: f(x) \geqslant u, x \in S\}$ are closed for each u); e.g., consider $f(t) \equiv \{1 \text{ if } 0 \leqslant t < 1 \text{ or } 1 < t \leqslant \infty, 0 \text{ if } t = 1\}$. Martos [25] called functions which satisfied (18) "explicitly quasiconcave" but he also assumed that f was continuous in which case (18) implies that f is also quasiconcave. Hanson [16] calls a function which satisfies (18) "functionally concave."

The following characterization of semistrict quasiconcavity is analogous to our third characterization of quasiconcavity.

THEOREM 7. *A continuous from above function f defined over S is semistrictly quasiconcave iff f satisfies the line segment minimum property* (1) *and* (19) *below:*

$$\begin{aligned} &x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, x^0 - \bar{t}v \in S, g(t) \\ &\quad \equiv f(x^0 + tv), g \text{ attains a local minimum at } 0 \text{ implies } g \\ &\quad \text{does not attain a one sided semistrict local minimum at } 0. \end{aligned} \tag{19}$$

COROLLARY 7.1. *A continuous function f defined over S is semistrictly quasiconcave iff f satisfies* (19).

COROLLARY 7.2. *A directionally differentiable function f defined over S is semistrictly quasiconcave iff*

$$\begin{aligned} &x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, x^0 - \bar{t}v \in S, D_v f(x^0) = 0 \\ &\quad \text{implies } g(t) \equiv f(x^0 + tv) \text{ defined for } x^0 + tv \in S \text{ does not} \\ &\quad \text{attain a local minimum at } t = 0 \text{ that is also a one-sided} \\ &\quad \text{semistrict local minimum.} \end{aligned} \tag{20}$$

COROLLARY 7.3. *A twice continuously differentiable function f defined over an open S is semistrictly quasiconcave iff*

$$x^0 \in S, \qquad v^T v = 1, \qquad v^T \nabla f(x^0) = 0$$

implies

$$\begin{aligned} &\text{(i) } v^T \nabla^2 f(x^0)\, v < 0 \text{ or (ii) } v^T \nabla^2 f(x^0)\, v = 0 \text{ and } g(t) \equiv f(x^0 + tv) \\ &\quad \text{does not attain a local minimum at } t = 0 \text{ that is also a} \\ &\quad \text{one sided semistrict local minimum.} \end{aligned} \tag{21}$$

4. STRICT QUASICONCAVITY

DEFINITION 6 (Ponstein [29, p. 118][5]). f is a *strictly quasiconcave* function defined over S iff

$$x^1 \in S, \qquad x^2 \in S, \qquad x^1 \neq x^2, \qquad f(x^1) \leqslant f(x^2), \qquad 0 < \lambda < 1$$

---

[5] Ponstein uses the term "unnamed concavity" in place of strict quasiconcavity while Thompson and Parke [35] use the term "X concavity." However, the term strict quasiconcavity seems to be consistent with the terminology used in economics; e.g. Katzner [21, p. 210] or Ginsberg [15, p. 598].

implies

$$f(x^1) < f(\lambda x^1 + (1-\lambda)x^2). \tag{22}$$

It is easy to see (cf. Ponstein [29, p. 118] and Thompson and Parke [35, p. 309]) that a strictly quasiconcave function is semistrictly quasiconcave and also quasiconcave. Moreover, the indifference sets (level curves) $I(u) \equiv \{x : f(x) = u, x \in S\}$ of a strictly quasiconcave function do not contain any straight line segments.

The following theorem is a slight extension of Theorem 6.

THEOREM 8. *If f is a strictly quasiconcave function defined over S and f attains a local maximum at $x^0 \in S$, then x^0 is the unique global maximizer of f over S.*

The following characterization of strict quasiconcavity is analogous to our third characterization of quasiconcavity.

THEOREM 9. *A function f defined over S is strictly quasiconcave iff f has the line segment minimum property* (1) *and* (23) *below:*

$$\begin{aligned} &x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, x^0 - \bar{t}v \in S \text{ implies } g(t) \\ &\quad \equiv f(x^0 + tv) \text{ does not attain a local minimum at } t = 0. \end{aligned} \tag{23}$$

COROLLARY 9.1. *A continuous* (*or continuous from below*) *function f defined over S is strictly quasiconcave iff f satisfies* (23).

COROLLARY 9.2. *A directionally differentiable function f defined over S is strictly quasiconcave iff the following condition holds:*

$$\begin{aligned} &x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, x^0 - \bar{t}v \in S, D_v f(x^0) = 0 \\ &\quad \text{implies } g(t) \equiv f(x^0 + tv) \text{ does not attain a local} \\ &\quad \text{minimum at } t = 0. \end{aligned} \tag{24}$$

COROLLARY 9.3. *A twice continuously differentiable function f defined over an open S is strictly quasiconcave iff*

$$\begin{aligned} &x^0 \in S, v^T v = 1, v^T \nabla f(x^0) = 0 \text{ implies (i) } v^T \nabla^2 f(x^0) v < 0 \\ &\quad \text{or (ii) } v^T \nabla^2 f(x^0) v = 0 \text{ and } g(t) \equiv f(x^0 + tv) \text{ does not} \\ &\quad \text{attain a local minimum at } t = 0. \end{aligned} \tag{25}$$

5. PSEUDOCONCAVITY

DEFINITION 7 (Mangasarian[6] [23, p. 281]). *A directionally differentiable function f defined over S is pseudoconcave iff f satisfies:*

$$x^0 \in S, v^Tv = 1, t > 0, D_v f(x^0) \leqslant 0 \text{ implies } f(x^0 + tv) \leqslant f(x^0). \tag{26}$$

Mangasarian [24, pp. 114–145] shows that pseudoconcave functions are also quasiconcave and semistrictly quasiconcave and have the following additional useful properties: (i) if x^0 is a local maximizer for the pseudoconcave function f, then it is also a global maximizer over S and (ii) if f is differentiable, S is an open convex set and $\nabla f(x^0) = 0_N$, then x^0 is a global maximizer for f. Ponstein [29] shows, however, that pseudoconcave functions are not necessarily strictly quasiconcave.

THEOREM 10. *Let f be a directionally differentiable function defined over S. Then f is pseudoconcave over S iff*

$$\begin{aligned} &x^0 \in S, v^Tv = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, D_v f(x^0) = 0 \textit{ implies} \\ &\quad g(t) \equiv f(x^0 + tv) \textit{ (defined for all } t \textit{ such that} \\ &\quad x^0 + tv \in S) \textit{ attains a local maximum at } t = 0. \end{aligned} \tag{27}$$

COROLLARY 10.1. *A twice continuously differentiable function f defined over an open S is pseudoconcave iff*

$$\begin{aligned} &x^0 \in S, v^Tv = 1, v^T\nabla f(x^0) = 0 \textit{ implies } \text{(i) } v^T\nabla^2 f(x^0)\, v < 0 \\ &\quad \textit{or } \text{(ii) } v^T\nabla^2 f(x^0)\, v = 0 \textit{ and } g(t) \equiv f(x^0 + tv) \textit{ attains} \\ &\quad \textit{a local maximum at } t = 0. \end{aligned} \tag{28}$$

Thus a one dimensional function having the property that all of its critial points are local maximums is pseudoconcave, and this fact can serve to characterize the class of pseudoconcave functions.

Additional conditions for pseudoconcavity in the twice differentiable case are given in Crouzeix and Ferland [6]. Finally, using a result due to Ferland [12, p. 300], it can be shown that the word "quasiconcave" in Theorem 5 above can be replaced by the word "pseudoconcave."

---

[6] Actually Mangasarian assumed that S is an *open* convex set and that f is differentiable, in which case (26) becomes: $x^0 \in S$, $x^1 \in S$, $(x^1 - x^0)^T \nabla f(x^0) \leqslant 0$ implies $f(x^1) \leqslant f(x^0)$. Thompson and Parke [35, p. 311] and Diewert [9] provide more general definitions of pseudoconcavity for the case where f is not necessarily differentiable.

6. STRICT PSEUDOCONCAVITY

DEFINITION 8 (Ponstein[7] [29, p. 117]). Let f be a directionally differentiable function defined over S. Then f is *strictly pseudoconcave* over S iff

$$x^0 \in S,\ v^Tv = 1,\ t > 0,\ x^0 + tv \in S,\ D_v f(x^0) \leqslant 0 \text{ implies } f(x^0 + tv) < f(x^0). \tag{29}$$

Comparison of Definitions 7 and 8 shows that strict pseudoconcavity implies pseudoconcavity. Ponstein [29, p. 118] shows that pseudoconcavity plus strict quasiconcavity implies strict pseudoconcavity.

THEOREM 11. *Let f be a directionally differentiable function defined over S. Then f is strictly pseudoconcave over S iff*

$$x^0 S,\ v^Tv = 1,\ \bar{t} > 0,\ x^0 + \bar{t}v \in S,\ D_v f(x^0) = 0 \text{ implies } g(t) \equiv f(x^0 + tv) \text{ attains a strict local maximum at } t = 0. \tag{30}$$

COROLLARY 11.1. *A twice continuously differentiable function f defined over an open S is strictly pseudoconcave iff*

$$x^0 \in S,\ v^Tv = 1,\ v^T\nabla f(x^0) = 0 \text{ implies (i) } v^T\nabla^2 f(x^0)\, v < 0 \text{ or (ii) } v^T\nabla^2 f(x^0)\, v = 0 \text{ and } g(t) \equiv f(x^0 + tv) \text{ attains a strict local maximum at } t = 0. \tag{31}$$

Comparison of properties (24) and (30) shows that a strictly pseudoconcave function is also strictly quasiconcave (a result which can be found in Ponstein [29]). Both of these types of functions share the following properties: (i) every local maximum is a unique global maximum and (ii) the indifference sets (level curves) $I(u) \equiv \{x: f(x) = u, x \in S\}$ do not contain any line segments. A strictly pseudoconcave function has the additional property *that $D_v f(x^0) = 0$ for every feasible direction v implies x^0 is the unique strict* global maximizer of f over S. Thus a strictly pseudoconcave function of one variable cannot have a point of inflection whereas a strictly quasiconcave function can.

---

[7] Ponstein's definition assumed that f had continuous first order partial derivatives over a convex set S and his definition was: $x^1 \in S$, $x^2 \in S$, $x^1 \neq x^2$, $f(x^2) \geqslant f(x^1)$ implies $(x^2 - x^1)^T \nabla f(x^1) > 0$. Definition 8 is a straightforward generalization of the contrapositive of Ponstein's definition.

7. STRONG QUASICONCAVITY OR STRONG PSEUDOCONCAVITY

DEFINITION 9. Let f be a directionally differentiable function defined over S. Then f is *strongly quasiconcave*[8] over S iff f is strictly quasiconcave over S and in addition has the following property:

$$x^0 \in S,\ v^Tv = 1,\ \bar{t} > 0,\ x^0 + \bar{t}v \in S,\ D_v f(x^0) = 0 \text{ implies there exist } \varepsilon \text{ and } \alpha \text{ such that } 0 < \varepsilon < \bar{t},\ 0 < \alpha \text{ and } f(x^0 + tv) < f(x^0) - \alpha t^2 \text{ for } t \in [0, \varepsilon]. \tag{32}$$

The above definition implies that a strongly quasiconcave function is a strictly quasiconcave function that has the additional property that if the directional derivative of the function is zero in a certain direction v at a point $x^0 \in S$ then the function decreases quadratically in a neighbourhood of x^0 along v.

Note that it is possible to characterize strong quasiconcavity using (32) and (19). However, it is simpler to characterize strongly quasiconcave functions using the concept of a strong local maximum

DEFINITION 10. A function of one variable g defined over a set S attains a *strong local maximum* at $t_0 \in S$ iff there exist $\alpha > 0$ and $\varepsilon > 0$ such that $g(t) \leqslant g(t_0) - \alpha(t - t_0)^2$ for $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ and $t \in S$.

THEOREM 12. *Let f be a directionally differentiable function defined over S. Then f is strongly quasiconcave over S iff f satisfies:*

$$x^0 \in S,\ v^Tv = 1,\ \bar{t} > 0,\ x^0 + \bar{t}v \in S,\ D_v f(x^0) = 0 \text{ implies } g(t) \equiv f(x^0 + tv) \text{ attains a strong local maximum at } t = 0. \tag{33}$$

COROLLARY 12.1. *A twice continuously differentiable function f defined over an open S is strongly quasiconcave iff*

$$x^0 \in S,\ v^Tv = 1,\ v^T\nabla f(x^0) = 0 \text{ implies } v^T\nabla^2 f(x^0)\, v < 0. \tag{34}$$

If $\nabla f(x^0) = 0_N$, then (34) is equivalent to negative definiteness of the Hessian matrix $\nabla^2 f(x^0)$ (recall conditions (9)), while if $\nabla f(x^0) \neq 0_N$, then (34) is equivalant to negative definiteness of $\nabla^2 f(x^0)$ in the subspace orthogonal to the gradient vector $\nabla f(x^0)$ (recall conditions (10)). Thus strong quasiconcavity in the twice differentiable case can be characterized by conditions (9) and (10), conditions which occur frequently in the economics literature.

---

[8] In a corollary to Theorem 12 below, we show that our definition of strong quasiconcavity is equivalent to the definition used by Newman [27, p. 307] and Ginsberg [15, p. 600], assuming that f is twice continuously differentiable and is defined over an open convex set. In place of the term "strong quasiconcavity" McFadden [26, p. 403] uses the term "differential strict quasiconcavity." Diewert [9] provides a definition of strong quasiconcavity that does not require the existence of directional derivatives.

Comparison of Thoerems 11 and 12 shows that strong quasiconcavity implies strict pseudoconcavity (the reverse implication is not true). Hence we could also call the class of strongly quasiconcave functions the class of strongly pseudoconcave functions.

8. CONCAVITY

DEFINITION 11. f is a concave[9] function over S iff

$$x^1 \in S, x^2 \in S, 0 \leqslant \lambda \leqslant 1 \text{ implies} f(\lambda x^1 + (1-\lambda)x^2) \geqslant f(x^1) + (1-\lambda) f(x^2). \tag{35}$$

Ponstein [29] shows that a differentiable concave function is necessarily quasiconcave, semistrictly quasiconcave and pseudoconcave, but it need not be strictly quasiconcave, strictly pseudoconcave or strongly quasiconcave.

THEOREM 13 (Fenchel [11, p. 81], Roberts and Varberg [32, p. 12]). *Let f be a directionally differentiable function defined over S. Then f is concave over S iff*

$$x^0 \in S, v^T v = 1, t > 0, x^0 + tv \in S \textit{ implies } f(x^0 + tv) \leqslant f(x^0) + tD_v f(x^0). \tag{36}$$

COROLLARY 13.1 (Mangasarian [24, p. 84]). *Let f be a differentiable function defined over an open S. Then f is concave over S iff*

$$x^0 \in S, x^1 \in S \textit{ implies} f(x^1) \leqslant f(x^0) + (x^1 - x^0)^T \nabla f(x^0). \tag{37}$$

Thus every linear approximation (tangent line) to a concave function is coincident with or lies above the function. Fenchel [11, pp. 87–89] and Mangasarian [24, p. 89] also show that if f is a twice continuously differentiable function defined over an open convex set S, then the following condition is necessary and sufficient for f to be concave.

$$x^0 \in S, v^T v = 1 \text{ implies } v^T \nabla^2 f(x^0)\, v \leqslant 0. \tag{38}$$

In other words, negative semidefiniteness of the Hessian matrix of f over S is necessary and sufficient for f to be concave over S.

---

[9] See the historical references in Fenchel [11] or Rockafellar [30].

The following two theorems provide alternative local characterizations of concavity, without necessarily assuming twice differentiability.

THEOREM 14. *f is concave over S iff*

$$x^0 \in S, v^T v = 1, t > 0, x^0 + tv \in S, a \in R^1, g(t) \equiv f(x^0 + tv) \text{ implies } h(t) \equiv g(t) - at \text{ is a quasiconcave function of one variable where } h(t) \text{ is defined for all } t \text{ such that } x^0 + tv \in S, \tag{39}$$

or equivalently (using Theorem 4) iff

$$x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, a \in R^1, h(t) \equiv f(x^0 + tv) - at \text{ (defined for all } t \in [0, \bar{t}]) \text{ implies } h \text{ has the line segment minimum property (1) and } h \text{ does not attain a semistrict local minimum at any } t_0 \in (0, \bar{t}). \tag{40}$$

THEOREM 15. *Let f be a directionally differentiable function defined over S. Then f is concave over S iff*

$$x^0 \in S, v^T v = 1, t > 0, x^0 + tv \in S, a \in R^1, g(t) \equiv f(x^0 + tv) \text{ implies } h(t) \equiv g(t) - at \text{ is a pseudoconcave function of one variable}, \tag{41}$$

or equivalently (using Theorem 10) iff

$$x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S, g(t) \equiv f(x^0 + tv), t_0 \in [0, \bar{t}], g'(t_0) \equiv a \text{ implies } h(t) \equiv g(t) - at \text{ attains a local maximum over } [0, \bar{t}] \text{ at } t_0. \tag{42}$$

COROLLARY 15.1. *Let f be a differentiable function defined over an open S. Then f is concave over S iff*

$$x^0 \in S, v^T v = 1 \text{ implies } h(t) \equiv f(x^0 + tv) - tv^T \nabla f(x^0) \text{ attains a local maximum at } t = 0. \tag{43}$$

COROLLARY 15.2. *Let f be a twice continuously differentiable function defined over an open S. Then f is concave over S iff*

$$x^0 \in S, v^T v = 1 \text{ implies (i) } v^T \nabla^2 f(x^0) v < 0 \text{ or (ii) } v^T \nabla^2 f(x^0) v = 0 \text{ and } h(t) \equiv f(x^0 + tv) - tv^T \nabla f(x^0) \text{ attains a local maximum at } t = 0. \tag{44}$$

The characterization of concavity given by Corollary 15.2 is equivalent to (38), but it is not quite as simple as (38). However, the characterization of concavity in the once differentiable case given by Corollary 15.1 above involves only the local properties of the function and hence will often be easier to check than the characterization given by Corollary 13.1.

9. STRICT CONCAVITY

DEFINITION 12 (Fenchel [11, p. 57]). f is *strictly concave* over S iff

$$x^1 \in S, x^2 \in S, x^1 \neq x^2, 0 < \lambda < 1 \text{ implies}$$
$$f(\lambda x^1 + (1-\lambda) x^2) > \lambda f(x^1) + (1-\lambda) f(x^2). \tag{45}$$

Ponstein [29] shows that a differentiable strictly concave function is quasiconcave, semistrictly quasiconcave, strictly quasiconcave, pseudoconcave and strictly pseudoconcave. However, a strictly concave function need not be strongly quasiconcave.

THEOREM 16 (Roberts and Varberg [32, p. 14]. *Let f be a directionally differentiable function defined over S. Then f is strictly concave over S iff*

$$x^0 \in S, v^T v = 1, t > 0, x^0 + tv \in S \textit{ implies}$$
$$f(x^0 + tv) < f(x^0) + tD_v f(x^0). \tag{46}$$

COROLLARY 16.1 (Ponstein [29, p. 115], Mangasarian [24, p. 87]). *Let f be a differentiable function defined over an open S. Then f is strictly concave over S iff*

$$x^0 \in S, x^1 \in S, x^0 \neq x^1 \textit{ implies}$$
$$f(x^1) < f(x^0) + (x^1 - x^0)^T \nabla f(x^0). \tag{47}$$

Thus every linear approximation (tangent line) to a strictly concave function lies above the function (except at the point of tangency). Fenchel [11, pp. 73–88] and Bernstein and Toupin [4, p. 72] show that if f is a twice continuously differentiable function defined over an open convex set S, then the following condition is necessary and sufficient for f to be strictly concave:

$$x^0 \in S, v^T v = 1, \bar{t} > 0, x^0 + \bar{t}v \in S \text{ implies}$$
$$\text{(i) } v^T \nabla^2 f(x^0) v \leqslant 0 \text{ and (ii) } v^T \nabla^2 f(x^0 + tv) v \neq 0$$
$$\textit{for almost all } t \in [0, \bar{t}]. \tag{48}$$

The following two theorems provide alternative local characterizations of strict concavity, without necessarily assuming twice differentiability.

THEOREM 17. *f is strictly concave over S iff*

$$\begin{aligned}&x^0 \in S,\ v^T v = 1,\ t > 0,\ x^0 + tv \in S,\ a \in R^1,\ g(t) \equiv f(x^0 + tv)\\ &\quad \textit{implies } h(t) \equiv g(t) - at \textit{ is a strictly quasiconcave}\\ &\quad \textit{function of one variable,} \qquad (49)\end{aligned}$$

or equivalently (*using Theorem* 9) *iff*

$$\begin{aligned}&x^0 \in S,\ v^T v = 1,\ \bar{t} > 0,\ x^0 + \bar{t}v \in S,\ a \in R^1,\ h(t) \equiv f(x^0 + tv) - at\\ &\quad \textit{implies } h \textit{ has the line segment minimum property } (1) \textit{ over}\\ &\quad [0, \bar{t}] \textit{ and } h \textit{ does not attain a local minium at any } t_0 \in (0, \bar{t}). \qquad (50)\end{aligned}$$

THEOREM 18. *Let f be a directionally differentiable function defined over S. Then f is strictly concave over S iff*

$$\begin{aligned}&x^0 \in S,\ v^T v = 1,\ t > 0,\ x^0 + tv \in S,\ a \in R^1,\ g(t) \equiv f(x^0 + tv) \textit{ implies}\\ &\quad h(t) \equiv g(t) - at \textit{ is a strictly pseudoconcave function,} \qquad (51)\end{aligned}$$

or equivalently (*using Theorem* 11) *iff*

$$\begin{aligned}&x^0 \in S,\ v^T v = 1,\ \bar{t} > 0,\ x^0 + \bar{t}v \in S,\ g(t) \equiv f(x^0 + tv),\\ &\quad t_0 \in [0, \bar{t}],\ g'(t_0) \equiv a \textit{ implies } h(t) \equiv g(t) - at\\ &\quad \textit{attains a strict local maximum over } [0, \bar{t}] \textit{ at } t_0. \qquad (52)\end{aligned}$$

COROLLARY 18.1. *Let f be a differentiable function defined over an open S. Then f is strictly concave over S iff*

$$\begin{aligned}&x^0 \in S,\ v^T v = 1 \textit{ implies } h(t) \equiv f(x^0 + tv) - tv^T \nabla f(x^0)\\ &\quad \textit{attains a strict local maximum at } t = 0. \qquad (53)\end{aligned}$$

COROLLARY 18.2. *Let f be a twice continuously differentiable function defined over an open S. Then f is strictly concave over S iff*

$$\begin{aligned}&x^0 \in S,\ v^T v = 1 \textit{ implies } (\mathrm{i})\ v^T \nabla^2 f(x^0)\, v < 0 \textit{ or } (\mathrm{ii})\ v^T \nabla^2 f(x^0)\, v = 0\\ &\quad \textit{and } h(t) \equiv f(x^0 + tv) - tv^T \nabla f(x^0) \textit{ attains a strict local max at}\\ &\quad t = 0. \qquad (54)\end{aligned}$$

We note that (54) is equivalent to (48).

10. STRONG CONCAVITY

DEFINITION 13 (Poljak [28, p. 73]. f is *strongly concave* over S iff there exists $a > 0$ such that for every $x^1 \in S$, $x^2 \in S$, $0 \leqslant \lambda \leqslant 1$,

$$\begin{aligned}&f(\lambda x^1 + (1 - \lambda)\, x^2)\\ &\quad \geqslant \lambda f(x^1) + (1 - \lambda) f(x^2) + \lambda(1 - \lambda)\, a (x^1 - x^2)^T (x^1 - x^2). \qquad (55)\end{aligned}$$

It is obvious from the above definition that a strongly concave function is strictly concave. The following theorems provide alternative characterizations of strong concavity.

THEOREM 19 (Rockafellar [31, p. 75]). *f is strongly concave over S iff there exists $\alpha > 0$ such that g defined by $g(x) \equiv f(x) + \alpha x^T x$ s is a concave function over S.*

THEOREM 20 (Poljak [28, p. 73]). *Let f be a differentiable function defined over an open S. Then f is strongly concave over S iff there exists an $\alpha > 0$ such that for every $x^1 \in S$, $x^2 \in S$,*

$$f(x^2) \leqslant f(x^1) + (x^2 - x^1)^T \nabla f(x^1) - \alpha(x^2 - x^1)^T (x^2 - x^1). \tag{56}$$

THEOREM 21 (Poljak [28, p. 73]). *Let f be a twice continuously differentiable function defined over an open S. Then f is strongly concave over S iff there exists an $\alpha > 0$ such that*

$$x^0 \in S,\ v^T v = 1 \text{ implies } v^T \nabla^2 f(x^0)\, v \leqslant -2\alpha. \tag{57}$$

Note that (57) implies the following condition:

$$x^0 \in S,\ v^T v = 1 \text{ implies } v^T \nabla^2 f(x^0)\, v < 0, \tag{58}$$

i.e., that the Hessian matrix of f be negative definite over S. Ginsberg [15, p. 597], Jorgenson and Lau [17, p. 40] and Lau [22, p. 135] use property (58) in order to define strong concavity in the twice differentiable case. We note that (58) is equivalent to (57) if f is a twice continuously differentiable function defined over a compact convex set S with a nonempty interior. Finally, note that (58) implies (54) and (48).

It can be shown that if a function is strongly concave, then it also satisfies each of the other eight types of quasiconcavity and concavity.

11. CONCLUSION

We have provided definitions and alternative characterizations of nine kinds of quasiconcavity and concavity that occur frequently in applied work in economics and management science.[10] In particular, we have provided characterizations involving the local behaviour of a function along line segments in the general case as well as in the once and twice differentiable case. These alternative characterizations should prove to be useful in applied work.

---

[10] See Diewert [8] for specific applications of the nine kinds of concavity in economics.

We conclude by noting that there are additional types of generalized concavity that we have not studied. For example, the class of quasiconcave functions that can be transformed by means of a monotonically increasing function of one variable into a concave function defines yet another type of quasiconcavity. This class of functions was studied by Fenchel [11] and more recently by Kannai [18, 19]. Avriel [2, pp. 160–172] and Zang [36] also study other classes of concave transformable functions and give extensive references to the literature.

APPENDIX: PROOFS OF THEOREMS

Proof of Theorem 3. The proof of this theorem follows readily from some monotonicity properties of quasiconcave functions developed by Stoer and Witzgall [34, p. 172].

Proof of Theorem 4. *Quasiconcavity implies* (1) *and* (2). Property (1) follows from Theorem 3. Hence we need only show that (1) and not (2) imply that f is not quasiconcave. This is straightforward.

Properties (1) *and* (2) *imply* f *is quasiconcave.* Let $x^0 \in S$, $x^1 \in S$, $x^0 \neq x^1$ and $0 < \lambda^* < 1$. *Suppose* that $f(\lambda^* x^0 + (1-\lambda^*)x^1) < \min\{f(x^0), f(x^1)\}$. Define $t_1 \equiv [(x^1 - x^0)^t (x^1 - x^0)]^{1/2} > 0$, $v \equiv (x^1 - x^0)/t_1$, and $g(t) \equiv f(x^0 + tv)$ for $t \in [0, t_1]$. By (1),

$$\begin{aligned}
\inf_t\{g(t)\colon t \in [0, t_1]\} &= \min_t\{g(t)\colon t \in [0, t_1]\} \\
&= \min_\lambda\{f(\lambda x^0 + (1-\lambda)x^1)\colon \lambda \in [0, 1]\} \\
&\leqslant f(\lambda^* x^0 + (1-\lambda^*)x^1) && \text{since } \lambda^* \text{ is feasible} \\
&< \min\{f(x^0), f(x^1)\} && \text{by our supposition} \\
&= \min\{g(0), g(t_1)\} && \text{by the definition of } g.
\end{aligned}$$

The above inequality shows that g attains a semistrict local minimum at some $t_0 \in (0, t_1)$ which contradicts (2). Thus our *supposition* is false and f is quasiconcave over S.

Proof of Theorem 5. A rather lengthy proof is available from the first author upon request.

Proof of Theorem 7. *Semistrict quasiconcavity implies* (1) *and* (19). Semistrict quasiconcavity implies quasiconcavity by a result of Karamardian [20] and hence by Theorem 3, (1) holds. Hence we need only show that (1) and not (19) imply that f is not semistrictly quasiconcave. This is straightforward.

Properties (1) *and* (19) *imply f is semistrictly quasiconcave.* Let $x^1 \in S$, $x^2 \in S$ and $f(x^1) < f(x^2)$. *Suppose* there exists λ^* such that $0 < \lambda^* < 1$ and $f(x^1) \geqslant f(\lambda^* x^1 + (1-\lambda^*) x^2)$. These inequalities imply that f attains a local minimum on the interior of the line segment joining x^1 to x^2 that is also a one sided semistrict local minimum. This contradicts (19). Thus our supposition is false and semistrict quasiconcavity of f follows.

Proof of Theorem 8. Suitably modify the proof of Theorem 6.

Proof of Theorem 9. *Strict quasiconcavity implies* (1) *and* (23). Since strict quasiconcavity implies quasiconcavity and quasiconcavity implies (1), we need only show that (1) and not (23) imply that f is not strictly quasiconcave. This is straightforward.

Properties (1) *and* (23) *imply f is strictly quasiconcave.* Let $x^1 \in S$, $x^2 \in S$, $x^1 \neq x^2$ and $f(x^1) \leqslant f(x^2)$. Define the set of minimizers of f over the closed interval joining x^1 and x^2 as M. If $M \equiv \{x^1\}$ or $M \equiv \{x^1, x^2\}$, then $f(x^1) < f(\lambda x^1 + (1-\lambda) x^2)$ for $0 < \lambda < 1$ and f is strictly quasiconcave.

Suppose $M \neq \{x^1\}$ and $M \neq \{x^1, x^2\}$. Then there exists a λ^0 such that $0 < \lambda^0 < 1$ and $f(\lambda^0 x^1 + (1-\lambda^0) x^2) \leqslant f(\lambda x^1 + (1-\lambda) x^2)$ for all λ such that $0 \leqslant \lambda \leqslant 1$. Thus f must attain a local minimum on the line segment joining x^1 and x^2 which contradicts (23). Thus our supposition is false and f is strictly quasiconcave.

Proof of Theorem 10. *Pseudoconcavity implies* (27). Obvious.

Property (27) *implies f is pseudoconcave.* Let $x^0 \in S$, $v^T v = 1$, $\bar{t} > 0$, $D_v f(x^0) \leqslant 0$.

Suppose $f(x^0 + \bar{t} v) > f(x^0)$. *Case* (i): $D_v f(x^0) < 0$. In this case, it is clear that $g(t) \equiv f(x^0 + tv)$ attains at least one local minimum over the open interval $(0, \bar{t})$. Let t_0 be the largest such local minimizer. Then $g'(t_0) = D_v f(x^0 + t_0 v) = 0$ but t_0 is not a local maximizer for g, which contradicts property (27). *Case* (ii): $D_v f(x^0) = 0$. By (27), $g(t) \equiv f(x^0 + tv)$ attains a local maximum at $t = 0$. Thus there exists $\varepsilon > 0$ such that $g(t) \leqslant g(0)$ for $0 \leqslant t \leqslant \varepsilon$. Since $g(\bar{t}) \equiv f(x^0 + \bar{t} v) > g(0)$ by our supposition, it can be seen that $g(t)$ must attain at least one local minimum over the open interval $(0, \bar{t})$. Let t_1 be the largest such local minimizer. Then $g'(t_1) = D_v f(x^0 + t_1 v) = 0$ but t_1 is not a local maximizer for g, which contradicts property (27). Thus in both cases, we have a contradiction and thus our supposition is false. Therefore $f(x^0 + \bar{t} v) \leqslant f(x^0)$ and pseudoconcavity of f follows.

Proof of Theorem 11. Similarly to the proof of Theorem 10.

Proof of Theorem 12. *Strong quasiconcavity implies* (33). Obvious.

Property (33) *implies f is strongly quasiconcave.* Property (33) implies

(24), and hence by Corollary 9.2, f is strictly quasiconcave. Using Definition 10, it is clear that (33) implies (32). Hence f is strongly quasiconcave by Definition 9.

Proof of Corollary 12.1. *Strong quasiconcavity implies* (34). Let $x^0 \in S$, $v^T v = 1$, and $D_v f(x^0) = 0$. By (32) and the openness of S, there exist $\alpha > 0$ and $\varepsilon > 0$ such that $g(t) \leqslant g(0) - \alpha t^2$ for $t \in [-\varepsilon, \varepsilon]$ where $g(t) \equiv f(x^0 + tv)$. By Taylor's Theorem, for $t \in (0, \varepsilon]$,

$$\begin{aligned} g(t) &= g(0) + g'(0)\,t + (1/2)\,g''(t^*)\,t^2 && \text{for some } t^* \in (0, t) \\ &= g(0) + g''(t^*)\,t^2/2 && \text{since } g'(0) = D_v f(x^0) = 0 \\ &\leqslant g(0) - \alpha t^2 && \text{since } g(t) \leqslant g(0) - \alpha t^2. \end{aligned}$$

The above inequality implies $g''(t^*) \leqslant -2\alpha$, where $0 < t^* < t$. Now let t tend to zero. Since f is twice continuously differentiable, so is g and we obtain the following inequality: $g''(0) = v^T \nabla^2 f(x^0)\, v \leqslant -2\alpha < 0$ and thus (34) is true.

Property (34) *implies f is strongly quasiconcave.* Let $x^0 \in S$, $v^T v = 1$, $\bar{t} > 0$., $x^0 + \bar{t} v \in S$, $D_v f(x^0) = v^T \nabla f(x^0) = 0$. By (23), we need only show $g(t) \equiv f(x^0 + tv)$ attains a strong local maximum at $t = 0$. Let $\delta > 0$ be small enough so that $x^0 + \delta v \in S$ and define α by $\max_t \{v^T \nabla^2 f(x^0 + tv)\, v : 0 \leqslant t \leqslant +\delta\} \equiv -2\alpha$. Property (34) implies that $\alpha > 0$. By Taylor's Theorem, for $t \in (0, \delta)$,

$$\begin{aligned} g(t) &= g(0) + g'(0)\,t + g''(t^*)\,t^2/2 && \text{for some } t^* \in (0, t) \\ &= g(0) + g''(t^*)\,t^2/2 && \text{since } g'(0) = D_v f(x^0) = 0 \\ &\leqslant g(0) - \alpha t^2 \end{aligned}$$

since $g''(t^*) \leqslant -2\alpha$ by the definition of α. Thus g attains a strong local maximum over the interval $[0, \delta]$ at $t = 0$.

Proof of Theorem 14. *f concave implies* (40). Since f is concave, so are g and h. But h concave implies that h is also quasiconcave (cf. Fenchel [11; 117]) and thus conditions (39) and (40) are satisfied.

Property (40) *implies that f is concave.* Suppose that f is not concave. Then there exist $x^0 \in S$, $v^T v = 1$, $\bar{t} > 0$, $0 < \lambda < 1$ such that $x^0 + \bar{t} v \in S$ and $f(\lambda x^0 + (1 - \lambda)(x^0 + \bar{t} v)) < \lambda f(x^0) + (1 - \lambda) f(x^0 + \bar{t} v)$. Define $g(t) \equiv f(x^0 + tv)$ for $0 \leqslant t \leqslant \bar{t}$ and define the average slope of the function g between 0 and $\bar{t}$ as $\alpha^* \equiv [g(\bar{t}) - g(0)]/[\bar{t} - 0]$. Define $h(t) \equiv g(t) - \alpha^* t$ for $t \in [0, \bar{t}]$. If h does not have the line segment minimum property, then h is not quasiconcave by Theorem 3 and (40) does not hold. If h does have the line segment minimum property, then using our definition of α^* and the inequality $g(\lambda 0 + (1 - \lambda)\,\bar{t}) < \lambda g(0) + (1 - \lambda)\, g(\bar{t})$, it can be seen that h

attains a semistrict local minimum for some $t_0 \in (0, \bar{t})$. Hence again (40) does not hold.

Proof of Theorem 15. *f concave implies* (42). Let $x^0 \in S$, $v^T v = 1$, $\bar{t} > 0$, $x^0 + \bar{t}v \in S$, $g(t) \equiv f(x^0 + tv)$, $t_0 \in [0, \bar{t}]$, $g'(t_0) \equiv \alpha$ and $h(t) \equiv g(t) - \alpha t$. Since f is concave, so are g and h. Thus by Theorem 13, $h(t) \leqslant h(t_0) + h'(t_0)(t - t_0) = h(t_0) + 0(t - t_0) = h(t_0)$ for all $t \in [0, \bar{t}]$. Thus h attains a local maximum in $[0, \bar{t}]$ at t_0.

Property (42) *implies f is concave.* Let $x^0 \in S$, $v^T v = 1$, $\bar{t} > 0$, $x^0 + \bar{t}v \in S$, $g(t) \equiv f(x^0 + tv)$, $\alpha \equiv g'(0) = D_v f(x^0)$ and $h(t) \equiv g(t) - \alpha t$ for $t \in [0, \bar{t}]$. Note that $h'(0) = 0$. Since (42) is equivalent to (41) h is pseudoconcave over $[0, \bar{t}]$. Thus by (26), $h(\bar{t}) \leqslant h(0)$ since $h'(0) = 0$ or $g(\bar{t}) - \alpha\bar{t} \leqslant g(0) - \alpha 0$ by the definition of h, or $g(\bar{t}) \leqslant g(0) + g'(0)\,\bar{t}$ by the definition of α, or $f(x^0 + \bar{t}v) \leqslant f(x^0) + \bar{t}D_v f(x^0)$ by the definition of g. The last inequality implies f is concave by Theorem 13.

Proof of Theorem 17. Similar to the proof of Theorem 14.

Proof of Theorem 18. Similar to the proof of Theorem 15.

Proof of Theorem 19. If g defined by $g(x) \equiv f(x) + \alpha x^T x$ is concave over S for some $\alpha > 0$, then for every $x^1 \in S$, $x^2 \in S$, $0 \leqslant \lambda \leqslant 1$, we have

$$\begin{aligned} &f(\lambda x^1 + (1-\lambda)x^2) + \alpha[\lambda x^1 + (1-\lambda)x^2]^T[\lambda x^1 + (1-\lambda)x^2] \\ &\quad \geqslant \lambda f(x^1) + (1-\lambda)f(x^2) + \lambda\alpha x^{1T}x^1 + (1-\lambda)\alpha x^{2T}x^2. \end{aligned}$$

A straightforward calculation shows that the above inequality is equivalent to the inequality (55).

REFERENCES

1. K. J. Arrow and A. C. Enthoven, Quasiconcave programming, *Econometrica* **29** (1961), 779–800.
2. M. Avriel, "Nonlinear Programming: Analysis and Methods," Prentice–Hall, Englewood Cliffs, N.J., 1976.
3. R. Bellman, "Introduction to Matrix Analysis," McGraw–Hill, New York, 1960.
4. B. Bernstein and R. A. Toupin, Some properties of the Hessian matrix of a strictly convex function, *J. Reine und Angew. Mathe.* **210** (1962), 65–72.
5. J.-P. Crouzeix, "Contributions a l'étude des fonctions quasiconvexes," Ph.D. Dissertation, *Université de Clermont, France, October, 1977.*
6. J.-P. Crouzeix and J. A. Ferland, Criteria for quasi-convexity and pseudo-convexity: relationships and comparisons, paper presented at the NATO Advanced Study Institute, Vancouver, Canada, August, 4–15, 1980.
7. G. Debreu, Definite and semidefinite quadratic forms, *Econometrica* **20** (1952), 195–300.
8. W. E. Diewert, Generalized concavity and economics, paper presented at the NATO Advanced Study Institute, Vancover, Canada, August 4–15, 1980.

9. W. E. DIEWERT, Alternative characterizations of six kinds of quasiconcavity in the nondifferentiable case with applications to nonsmooth programming, paper presented at the NATO Advanced Study Institute, Vancouver, Canada, August 4–15, 1980.
10. W. E. DIEWERT AND A. D. WOODLAND, Frank Knight's theorem in linear programming revisited, *Econometrica* **45** (1977), 375–398.
11. W. FENCHEL, Convex cones, sets and functions, Lecture Notes, Department of Mathematics, Princeton University, 1953.
12. J. A. FERLAND, Mathematical programming problems with quasi-convex objective functions, *Math. Program.* **3** (1972), 296–301.
13. J. A. FERLAND, Matrix theoretic criteria for the quasi-convexity of twice continuously differentiable functions, paper presented at the NATO Advanced Study Institute, Vancouver, Canada, August 4–15, 1980.
14. R. M. ELKIN, "Convergence Theorems for Gauss–Seidel and Other Minimization Algorithms," Ph.D. Dissertation, University of Maryland, Colle Park, Maryland, 1968.
15. W. GINSBERG, Concavity and quasiconcavity in economics, *J. Econ. Theory* **6** (1973), 596–605.
16. M. A. HANSON, Bounds for functionally convex optimal control problems, *J. Math. Anal. Appl.* **8** (1964), 84–89.
17. D. W. JORGENSON AND L. J. LAU, Duality and differentiability in production, *J. Econ. Theory* **9** (1974), 23–42.
18. Y. KANNAI, Concavifiability and constructions of concave utility functions, *J. Math. Econ.* **4** (1977), 1–56.
19. Y. KANNAI, Concave utility functions-existence, constructions and cardinality, paper presented at the NATO Advanced Study Institute, Vancouver, Canada, August 4–15, 1980.
20. S. KARAMARDIAN, Duality in mathematical programming, *J. Math. Anal. Appl.* **20** (1967), 344–358.
21. D. W. KATZNER, "Static Demand Theory," MacMillan, New York, 1970.
22. L. J. LAU, Applications of profit functions, *in* "Production Economics: A Dual Approach to Theory and Applications" (M. Fuss and D. McFadden, Eds.), Vol. 1, pp. 133–216, North-Holland, Amsterdam, 1978.
23. O. L. MANGASARIAN, Pseudo-convex functions, *SIAM J. Control* **3** (1965), 281–29.
24. O. L. MANGASARIAN, "Nonlinear Programming," McGraw–Hill, New York, 1969.
25. B. MARTOS, The direct power of adjacent vertex programming methods, *Manage. Sci.* **12** (1965), 241–252.
26. D. MCFADDEN, Convex analysis, *in* "Production Economics: A Dual Approach to Theory and Applications," Vol. 1 (M. Fuss and D. McFadden, Eds.), pp. 383–408, North-Holland, Amsterdam, 1978.
27. P. NEUMAN, Some properties of concave functions, *J. Econ. Theory* **1** (1969), 291–314.
28. B. T. POLJAK, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Dokl. Akad. Nauk C.C.P.* **166** (1966), 287-290. (English translation in *Soviet Math.* **7** (1966), 72–75.)
29. J. PONSTEIN, Seven kinds of convexity, *SIAM Rev.* **9** (1967), 115–119.
30. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, N.J., 1970.
31. R. T. ROCKAFELLAR, Saddle points of Hamiltonian systems in convex Lagrange problems having a nonzero discount rate, *J. Econ. Theory* **12** (1976), 71–113.
32. A. W. ROBERTS AND D. E. VARBERG, "Convex Functions," Academic Press, New York, 1973.
33. P. A. SAMUELSON, "Foundations of Economic Analysis," Harvard Univ. Press, Cambridge, Mass., 1947.

34. J. STOER AND C. WITZGALL, "Convexity and Optimization in Finite Dimensions I," Springer-Verlag, Berlin, 1970.
35. W. A. THOMPSON JR. AND D. W. PARKE, Some properties of generalized concave functions, *Operations Res.* **21**(1973), 305–313.
36. I. ZANG, Concavifiability of C^2 functions: A unified exposition, paper presented at the NATO Advanced Study Institute, Vancouver, Canada, August 4–15, 1980.

JOURNAL OF ECONOMIC THEORY 25, 421–430 (1981)

A Hybrid Fiat–Commodity Monetary System*

NEIL WALLACE

University of Minnesota, Minneapolis, Minnesota 55455

Received September 8, 1980

1. INTRODUCTION

In this paper I describe a "monetary" system in which backing is provided for the government's liabilities by way of contingent resort to taxes. The system is studied within the context of a one-good, pure exchange model of two-period-lived overlapping generations in which, aside from various uniform boundedness assumptions, considerable diversity is allowed both within and across generations. Two results are established: (i) the existence of at least one perfect foresight competitive equilibrium, and (ii) the Pareto optimality of *any* such equilibrium.

The (monetary) system studied has some of the features of a commodity money system with a large seignorage spread between bid and ask prices. One crucial feature of the system is that the government, by taxing, stands ready at any time to convert its outstanding liabilities into the commodity at an announced price. In an important sense, the hybrid scheme is a mixture of a fixed supply, fiat scheme and a pure tax-transfer scheme (a social security scheme); it makes some use of market pricing of net government indebtedness and some use of taxes and transfers. Its one advantage vis-a-vis the fixed supply, fiat scheme is that *any* equilibrium under the hybrid scheme is Pareto optimal. For the fixed supply, fiat scheme one can at best demonstrate that there exists an optimal equilibrium; it is well known that there can exist nonoptimal equilibria under that scheme. The presumed advantage of the hybrid scheme vis-a-vis the pure tax-transfer scheme is its lesser use of taxes and transfers. I say presumed because the framework I set out is not rich enough to imply that less use of taxes and transfers is preferable to more use of taxes and transfers.

---

* Support for this research was provided by the Federal Reserve Bank of Minneapolis and by the National Science Foundation under Grant SOC 77-22743 to the University of Minesota. I am indebted to John Bryant and Preston Miller of the Federal Reserve Bank of Minneapolis, to my university colleagues, M. K. Richter and William Thomson, and to an anonymous referee for extremely helpful comments on an earlier draft. I alone, however, am responsible for any errors.

0022-0531/81/060421-10$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

The rest of this paper is organized as follows. In Section 2, I describe the economy studied. In Section 3, I describe the conditions for perfect foresight competitive equilibrium, establish existence of such an equilibrium, and—using a Balasko–Shell [1] optimality criterion—prove that any such equilibrium is Pareto optimal. Two examples are provided in Section 4 and concluding remarks in Section 5.

2. ECONOMY

2.1. *Endowments and Preferences*

Time is discrete and takes on the values $t = 1, 2, \ldots$. At each date t, $t \geqslant 1$, a new generation, generation t, of size $N(t)$ appears and is present in the economy at t and $t+1$. The $N(0)$ members of generation 0 are present only at $t = 1$. There is a single, nonproduced good in the economy at each date. Society's endowment of this good, denoted $W(t)$, is bounded; i.e., $W(t) \leqslant W$ for all $t \geqslant 1$.

Let $w_t^j(t) \geqslant 0$ be the endowment of time t good of member j of generation i and let $W_t(t) = \sum_j w_t^j(t)$ and $W_t(t+1) = \sum_j w_t^j(t+1)$, where summation over j is summation over the members of generation t. Here $W_t(t)(W_t(t+1))$ is the total amount of time t $(t+1)$ good owned by members of generation t. I assume $W_{t-1}(t) + W_t(t) = W(t)$, $W_t(t) \geqslant W_1 > 0$ and $W_t(t+1) \geqslant W_2 > 0$ for all $t \geqslant 1$; that is, individual endowments of time t good for members of generations $t-1$ and t exhaust society's endowment, $W(t)$, and the aggregate endowment of generation t of both time t and time $t+1$ goods is bounded away from zero. I also assume that individual endowments are uniformly bounded away from zero; i.e., there exists $w > 0$ such that $(w_t^j(t), w_t^j(t+1)) \geqslant (w, w)$ for all j and $t \geqslant 1$.

As regards preferences, each member of generation 0 maximizes consumption of time 1 good, while each member j of generation t, $t > 0$, has a twice differentiable utility function $u_t^j(\cdot, \cdot)$ that is defined on positive consumption of time t and time $t+1$ goods, respectively. The function u_t^j has positive first-order partial derivatives and displays a diminishing marginal rate of substitution. Moreover, if (x, y) is such that $(x, y) \leqslant (W, W)$ and $u_t^j(x, y) \geqslant u_t^j(w_t^j(t), w_t^j(t+1)/2)$ for some j and $t > 0$, then it is assumed (a) that $x > 0$ and $y > 0$, (b) that there exist scalars $h > 0$ and $H > 0$ (not dependent of j or t) such that $h \leqslant u_{t1}^j(x, y)/u_{t2}^j(x, y) \leqslant H$ (i.e., bounded marginal rates of substitution), and (c) that there exists another pair of scalars (not dependent of j or t) that provide a uniform and positive lower bound and a uniform and positive upper bound on the Gaussian curvature of any indifference curve of u_t^j at any such (x, y). (See Balasko and Shell [1] for a discussion of the role of these uniformity assumptions on curvatures.)

2.2. *The "Monetary" System*

The policy studied is described in terms of three parameters, $(\underline{p}, \bar{p})$ and M, and a supporting tax-transfer scheme, where $\underline{p}$ and $\bar{p}$ are positive prices of money—somewhat like bid and ask prices, respectively—and M is some positive number of units of money. At any date t, the government's supply curve (correspondence) of money is as shown in Fig. 1; $M(t)$ is the outstanding stock of money held by the public from t to $t+1$ and $p(t)$ is the market price of a unit of money at t in units of time t goods (the inverse of the price level at t). I assume that this scheme is set up at $t = 1$ and that $M(0) = 0$. The scheme is called a hybrid fiat-commodity scheme because with $\underline{p} = 0$ and $\bar{p}$ sufficiently large it is a fiat scheme with a fixed supply of fiat money M, while with $\underline{p} = \bar{p} > 0$ it is a commodity scheme with no seignorage spread.

To ensure feasibility of the hybrid scheme, I assume that the parameters $(\underline{p}, \bar{p})$ and M satisfy $M > 0$, $0 < \underline{p}M \leqslant W_2/2$, and $W \leqslant \bar{p}M$. Since $p(t)\,M(t)$, expenditure on money by the young at t, is bounded above by W, the last inequality implies $p(t)\,M(t) < \bar{p}M$ for all t. It follows that no money is ever purchased from the government at $\bar{p}$, and, therefore, that $M(t) \leqslant M$ for all t.

At time t, the total of goods handed out by the government under this scheme is $p(t)[M(t-1) - M(t)]$, which I denote by $\tau(t)$. Note that if

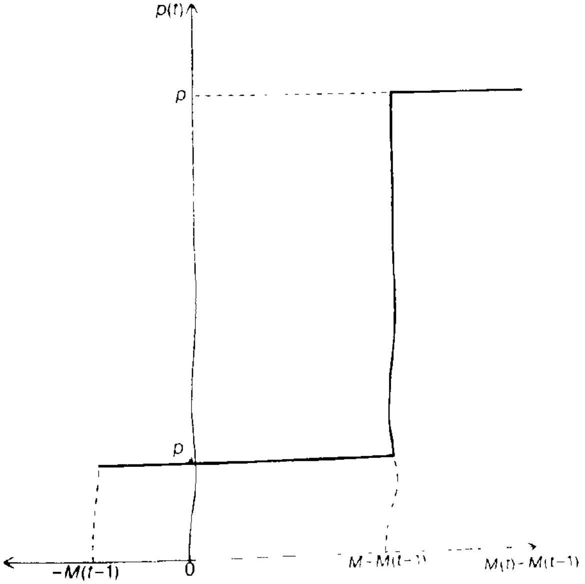


FIG. 1. Government supply of money at time t under the hybrid scheme.

$\tau(t) > 0$, then $M(t) < M$ and $p(t) = \underline{p}$. If follows that $\tau(t) \leqslant \underline{p}M$. I assume that $\tau(t)$ is financed by lump-sum taxes as follows.

All taxes (and transfers) are levied on the old. Thus, let $\tau_t^j(t+1)$ be the tax payable at $t+1$ by member j of generation t. I assume that $\tau_t^j(t+1) \leqslant w_t^j(t+1)/2$, that $\tau_t^j(t+1)$ is a continuous function (possibly dependent on j and t) of $\tau(t+1)$ and, of course, that $\sum_j \tau_t^j(t+1) = \tau(t+1)$. That the first and third of these assumptions can be satisfied follows from $\tau(t+1) \leqslant \underline{p}M \leqslant W_2/2 \leqslant \sum_j w_t^j(t+1)/2$. The second assumption, continuity, is important for the existence-of-equilibrium proof. All three assumptions are satisfied by the following "proportional" tax scheme: $\tau_t^j(t+1) = w_t^j(t+1)\,\tau(t+1)/W_t(t+1)$.

3. Perfect Foresight, Competitive Equilibrium: Existence and Optimality

From now on, unless otherwise noted, I will use the term equilibrium to refer to a perfect foresight competitive equilibrium.

I find it convenient to state equilibrium conditions in terms of an aggregate saving function for members of generation t. Thus, for member j of generation t, let $s_t^j[r(t), \tau_t^j(t+1)]$ be the unique and continuous solution for $w_t^j(t) - c_t^j(t)$ to the problem: choose $(c_t^j(t), c_t^j(t+1))$ to maximize $u_t^j(\cdot,\cdot)$ subject to

$$c_t^j(t) + c_t^j(t+1)/r(t) \leqslant w_t^j(t) + [w_t^j(t+1) - \tau_t^j(t+1)]/r(t). \tag{1}$$

Then, let

$$S_t[r(t), \tau(t+1)] = \sum_j s_t^j[r(t), \tau_t^j(t+1)].$$

So defined, the function S_t, the aggregate saving function for generation t, is well defined and continuous in each of its arguments.

The first result we state is one that relates equilibrium values of $r(t)$ to $p(t)$ and $p(t+1)$, where, recall, $p(t)$ is the price at t in terms of time t good of a unit of money. In equilibrium, $p(t+1) \geqslant \underline{p}$ and

$$\begin{aligned} r(t) &= p(t+1)/p(t) \qquad \text{if} \quad M(t) > 0 \\ &\geqslant p(t+1)/p(t) \qquad \text{if} \quad M(t) = 0. \end{aligned} \tag{2}$$

The first line of (2) is an arbitrage condition. It says that if money is held *from t to $t+1$, then the terms of trade between time t and time $t+1$ goods (the time t interest rate) implied by the values of money at t and $t+1$ must* be the market interest rate. As for the second line, it gives a lower bound on the market interest rate if no money is purchased by the young at t. In particular, it implies that if $M(t) = 0$, then $r(t) \geqslant 1$ because money is available to be purchased at $p(t) = \underline{p}$ and can be sold subsequently at $p(t+1) \geqslant \underline{p}$.

Next, we express $p(t+1)/p(t)$ in terms of real money holdings. Let $q(t) \equiv p(t)\,M(t)$ (real money holdings held from t to $t+1$ by generation t), let $q \equiv \underline{p}M$ and let R, a function on $[0, W] \times [0, W]$ be defined by

$$\begin{aligned} R(x, y) &= 1 && \text{if} \quad (0, 0) \leqslant (x, y) \leqslant (q, q) \\ &= y/x && \text{if} \quad (q, q) \leqslant (x, y) \leqslant (W, W) \\ &= y/q && \text{if} \quad (0, q) \leqslant (x, y) \leqslant (q, W) \\ &= q/x && \text{if} \quad (q, 0) \leqslant (x, y) \leqslant (W, q). \end{aligned} \tag{3}$$

So defined, R is continuous and, since the government supply function implies $[M(t) - M][p(t) - \underline{p}] = 0$, any equilibrium satisfies $p(t+1)/p(t) = R[q(t), q(t+1)]$.

As a last preliminary, we express $\tau(t+1)$ in terms of $q(t)$, $q(t+1)$, and $r(t)$. First, from the definition of $q(t)$, $\tau(t+1) = q(t)\,p(t+1)/p(t) - q(t+1)$. Then, since in equilibrium $p(t+1)/p(t) = r(t)$ except when $M(t)$ and, hence, $q(t)$ is zero, in equilibrium $\tau(t+1) = q(t)\,r(t) - q(t+1)$.

Assembling these results, we have the following

DEFINITION. An equilibrium is a nonnegative $q(t)$ sequence and an $r(t)$ sequence that for all $t \geqslant 1$ satisfy

$$S_t[r(t), q(t)\,r(t) - q(t+1)] = q(t), \tag{4}$$

$$r(t) - R[q(t), q(t+1)] \geqslant 0, \tag{5}$$

$$\{r(t) - R[q(t), q(t+1)]\}\, q(t) = 0. \tag{6}$$

Equation (4) says that aggregate saving must equal the value of money, while conditions (5) and (6) relate the market rate of interest to $q(t)$ and $q(t+1)$ as implied by (2) and the hybrid "monetary" system.

We now state and prove our two results.

PROPOSITION 1. *An equilibrium exists.*

Proof. The first step is to show that for any $q(t+1) \in [0, W]$, there exists at least one pair $[q(t), r(t)] \in [0, W] \times (0, \infty)$ that satisfies (4)–(6).

Let $f_t(x, y) = S_t[R(x, y), xR(x, y) - y]$. So defined, for any fixed $y \in [0, W]$, $f_t(x, y)$ is defined and continuous for all $x \in [0, W]$ and $f_t(x, y) < W$. There are two cases to consider. If $f_t(0, y) \geqslant 0$, then there exists at least one value of $x \in [0\ W]$ with $f_t(x, y) = x$. At any such x and

$r(t) = R(x, y)$, (4)–(6) are satisfied. If $f_t(0, y) < 0$ (desired saving is negative at $r(t) = R(0, y)$), then, by the continuity of S_t and by our endowment and preference assumptions, there exists some $r(t) > R(0, y)$ such that $S_t[r(t), -y] = 0$. This value of $r(t)$ and $q(t) = 0$ satisfy (4)–(6).

This establishes the existence of a mapping G_t that associates with each $q(t+1) \in [0, W]$ a nonempty subset of $[0, W] \times (0, \infty)$ with the property that $q(t+1)$ and $[q(t), r(t)] \in G_t(q(t+1))$ satisfy (4)–(6). Let us denote by $\gamma_t(q(t+1))$ the projection of $G_t(q(t+1))$ on $[0, W]$. It follows from the continuity of $f_t(x, y) - x$ in (x, y) and from the continuity of $S_t[r(t), -y]$ in its arguments that γ_t maps nonempty compact subsets of $[0, W]$ into nonempty compact subsets of $[0, W]$.

Now let $I = [0, W]$ and denote by $\gamma_t(B)$ the range of γ_t on the set B. Also, let $\{Q(t)\}$, a sequence of sets, each of which is an infinite product of intervals, be defined by

$$
\begin{aligned}
Q_0 &= \chi_1^\infty I,\\
Q_1 &= \gamma_1(I) \times \chi_2^\infty I,\\
Q_2 &= \gamma_1(\gamma_2(I)) \times \gamma_2(I) \times \chi_3^\infty I,\\
&\vdots\\
Q_t &= \gamma_1(\gamma_2(\cdots \gamma_t(I)) \cdots) \times \gamma_2(\cdots (\gamma_t(I))) \cdots) \times \cdots \times \gamma_t(I) \times \chi_{t+1}^\infty I,\\
&\vdots
\end{aligned}
$$

So defined, each Q_t is an infinite product of compact sets and, therefore, is compact. Moreover, by the property of γ_t established above, each Q_t is nonempty. Then since $Q_t \supset Q_{t+1}$ for all $t \geqslant 0$, it follows that $Q \equiv \bigcap_0^\infty Q_t$ is nonempty and that any element $\{q(t)\} \in Q$ satisfies $q(t) \in \gamma_t(q(t+1))$ for all $t \geqslant 1$. It is obvious from the first step in the proof that we can asociate with any such $q(t)$ sequence an $r(t)$ sequence such that both satisfy (4)–(6) for all $t \geqslant 1$. ∎

Proposition 2. *Any proposition 1 equilibrium is Pareto optimal.*

Proof. It follows from (2) that for any equilibrium and any $t \geqslant 1$, the product $\prod_1^t r(i) \geqslant \prod_1^t p(i+1)/p(i) = p(t+1)/p(1) \geqslant \underline{p}/\bar{p}$. Since the inverse of $\prod_1^t r(i)$ is the price of time $t+1$ good in units of time 1 good, it follows that the latter is bounded from above. As Balasko and Shell show (see Section 5 of their paper), this implies optimality under our assumptions. ∎

Note that the proof of Proposition 2 uses the perfect foresight aspect of the equilibrium. Optimality cannot be established using only the bounds on the price of money; namely, $\underline{p} \leqslant p(t) \leqslant \bar{p}$ for all t. When $M(t) > 0$, such bounds imply only $r(t) \geqslant \underline{p}/\bar{p}$, which allows for convergence of $\prod_1^t r(i)$ to zero. The crucial feature of a perfect foresight sequence is that if $r(t)$ is small because $p(t+1)$ is small, then $r(t+1)$ tends to be not so small because $p(t+1)$ is the denominator of $r(t+1)$. It would certainly seem, however, that if we had a notion of "almost" perfect foresight equilibrium, then optimality could also be established for such an equilibrium.

4. TWO EXAMPLES

Two very different examples are presented in this section. The first example is of the kind that motivated the hybrid scheme. The second shows how interventionist the hybrid scheme can be.

The first example is of an economy with many equilibria under a fixed supply, fiat scheme, many that are nonoptimal and one that is optimal. By Proposition 2, the nonoptimal ones cannot be equilibria under the hybrid scheme. As we will see, the optimal one is an equilibrium under the hybrid scheme if $\underline{p}M$ is small enough. The second example is of an economy with a unique equilibrium under a fixed supply, fiat scheme, an equilibrium which is optimal and is such that money does not have value. This equilibrium is not an equilibrium under the hybrid scheme.

EXAMPLE 1. $N(t)=1$, $u_t(c_1, c_2)=c_1c_2$ and $(w_t(t), w_t(t+1))=(w_1, w_2)$, $w_1 > w_2$.

Under the fixed supply, fiat scheme with say, $M=1$, the utility maximizing condition implies that any nonnegative $p(t)$ sequence satisfying $p(t)=w_1\,p(t+1)/[2p(t+1)+w_2]$ (shown in Fig. 2) and $p(t) < w_1$ is an equilibrium. Thus, as is evident from Fig. 2, any $p(1) \in [0, (w_1-w_2)/2)$ can be the first term of a nonnegative $p(t)$ sequence with $p(t)$ converging to zero that is an equilibrium. It is well known that all such equilibria are nonoptimal; for any such equilibrium, $r(t)$ approaches $w_2/w_1 < 1$. The only other equilibrium under the fixed supply, fiat scheme is $p(t)=(w_1-w_2)/2$ for all $t \geqslant 1$. It can be shown that this is the unique equilibrium under the hybrid scheme if $\underline{p}M < (w_1-w_2)/2$, an equilibrium in which, aside from a transfer to the current or initial old, no taxes or transfers are ever levied.

EXAMPLE 2. $N(t)=1$, $u_t(c_1, c_2)=c_1c_2$

$$\begin{aligned}(w_t(t), w_t(t+1)) &= (3/4, 1/4): \qquad t=1, 3, 5, \ldots\\ &= (1/8, 7/8): \qquad t=2, 4, 6, \ldots\end{aligned}$$

The unique equilibrium under the fixed supply, fiat scheme is autarkic and given by $r(t)=1/3$ for t odd and $r(t)=7$ for t even. Since $\prod_1^t r(i) \geqslant 1/3$ for all t, this is an optimal equilibrium.

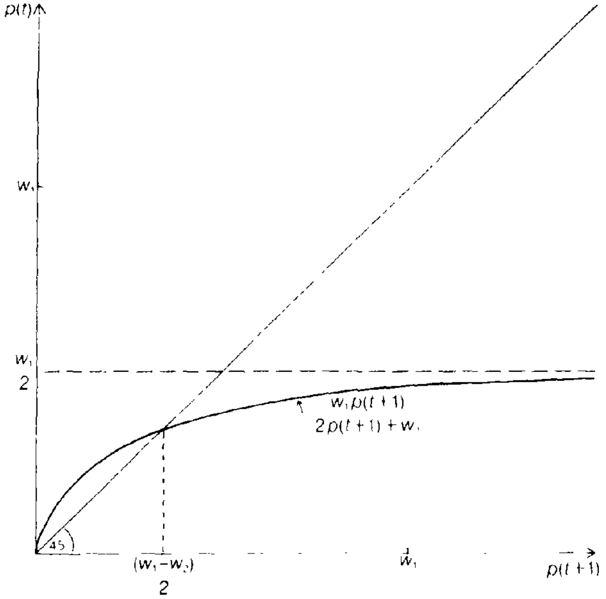


FIG. 2. Equilibrium difference equation for Example 1 under fixed supply, fiat money with $M = 1$.

It is easy to see that this autarkic equilibrium is not an equilibrium under the hybrid scheme. In order that it be one, $q(t) = 0$ for all $t \geqslant 1$ would have to be consistent with utility maximizing choices given the hybrid rates of return. But $q(t) = 0$ implies $r(t) \geqslant 1$ under the hybrid scheme, and in odd periods $q(t) = 0$ is not utility maximizing at any $r(t) \geqslant 1$.

If $M = 1$ and $\underline{p} \leqslant 1/8$, then the unique equilibrium under the hybrid scheme is $(M(t), p(t)) = (1, 3\underline{p}/(1 + 4\underline{p}))$ for t odd and $(M(t), p(t)) = (0, \underline{p})$ for t even. To verify that this is an equilibrium, one checks that $p(t) \geqslant \underline{p}$, that $3/4 - c_1 = 3\underline{p}/(1 + 4\underline{p})$ maximizes $c_1 c_2$ subject to the t-odd version of (1) and $3/4 - c_1 \geqslant 0$, and that $1/8 - c_1 = 0$ maximizes the same utility function subject to the t-even version of (1) and $1/8 - c_1 \geqslant 0$. The equilibrium lifetime allocations are displayed in Fig. 3. (Verifying uniqueness in either this or the last example is somewhat tedious. For that reason and because the proofs depend on the special features of the examples, they are not included.)

Note that the hybrid scheme equilibria for Examples 1 and 2 look very different. In the first example, a casual observer of the hybrid scheme equilibrium could mistakenly infer that, because no money is ever turned in, the cash-in price $\underline{p}$ plays no role. In the second example, the cash-in price is an obvious determinant of the equilibrium consumption allocation.

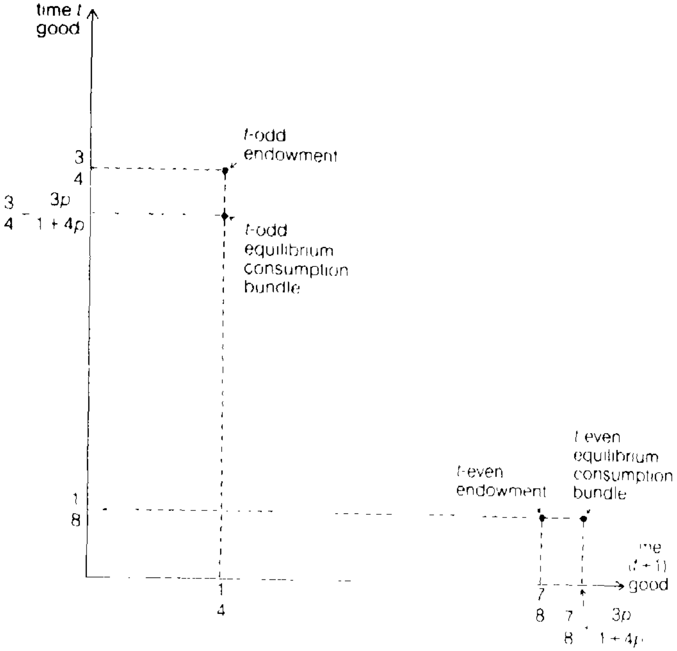


FIG. 3. The hybrid equilibrium for Example 2.

5. CONCLUDING REMARKS

I have shown that in an appropriately bounded economy, there exists at least one perfect foresight competitive equilibrium under the hydrid scheme and that any such equilibrium is Pareto optimal. As was noted above, the second of these results is not true for the fixed supply fiat scheme, a scheme in which government policy consists of nothing more than the provision of a fixed stock of fiat money. But this advantage of the hybrid scheme does not come for nothing.

First, as we have seen, it may be necessary to levy taxes under the hybrid scheme. This is presumably a disadvantage of the hybrid scheme vis a vis the fixed supply, fiat scheme. Second, implementation of an optimal hybrid scheme requires knowledge of at least some features of the economy. Thus, for example, the hybrid scheme described does not guarantee optimality in an economy in which $W(t)$ grows exponentially; in such an economy, optimality is implied by a hybrid scheme in which p and $\bar{p}$ grow exponentially at the same rate as does $W(t)$.

For these and other reasons, this paper does not constitute a plea for adoption of a monetary system like the hybrid scheme. Its main contribution is the demonstration that it is possible to analyze a somewhat complicated monetary system in the context of a model that is true to the notion that assets are acquired only in order to accomplish intertemporal trades. Although economists have been debating the virtues of various monetary systems for a long time, most discussions are not true to that notion, and, partly as a consequence, leave unanswered basic questions concerning feasibility and the desirability of a governmental role. The analysis presented above begins to address such questions.

REFERENCES

1. Y. BALASKO AND K. SHELL, The overlapping-generations model. I. The case of pure exchange without money, *J. Econ. Theory* **23** (1980), 281–306.

JOURNAL OF ECONOMIC THEORY 25, 431–441 (1981)

Notes, Comments, and Letters to the Editor

A Class of Solutions to Bargaining Problems

WILLIAM THOMSON*

Department of Economics, Harvard University, Cambridge, Massachusetts 02138

Received October 30, 1978; revised August 21, 1981

A reformation of Nash's axiom of Independence of Irrelevant Alternatives is proposed, generalizing an idea due to Roth. It involves the concept of a *reference function* g, whose purpose is to summarize the main features of the bargaining problem and so facilitate the evaluation of the relative bargaining positions of the players. It is with respect to g that the Independence condition is reformulated. In conjunction with the other axioms used by Nash, it is shown that if g satisfies certain natural properties, the bargaining problem admits of a unique solution. Several examples of reference functions are shown to satisfy these conditions. An existence problem discovered by Roth for a particular choice of g is shown to be resolved if solutions are permitted to be multi-valued, the axioms being correspondingly reformulated. *J. Econ. Theory* December 1981, 25(3), pp. 431–441. Harvard University, Cambridge, Massachusetts 02138 *Journal of Economic Literature*. Classification Number: 025.

1. INTRODUCTION

In [7], Roth obtained a new solution to the classic bargaining problem through a reformulation of the axiom of Independence of Irrelevant Alternatives which appears in Nash's original treatment [4]. A more general version of this axiom, encompassing Nash's and Roth's forms as particular cases is proposed here, involving the introduction of a *reference function*, a function singling out, for each bargaining problem, a point of the utility space summarizing its essential features and facilitating the evaluation of the relative bargaining strength of the players. It is with respect to this reference point that the Independence axiom is reformulated. By choosing reference

---

* University of Minnesota and Harvard University. Partial support from NSF (Grant SES800-6482) is gratefully acknowledged, as well as the referee's numerous suggestions for improvements.

0022-0531/81/060431-11$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

functions satisfying two natural properties, the modified axiom is shown to yield, in conjunction with the other axioms used in the characterization of Nash's solution, a large class of new solutions.

An appealing example of reference function, suggested in particular by Roth, turns out to lead to a non-existence result, as he showed. This negative result does not remain if multi-valued solutions are permitted. Showing this requires that the concept of a solution be redefined and the axioms reformulated. The existence of a solution satisfying all the axioms for that choice of reference function is then established.

The paper is organized as follows. Section 2 contains some preliminaries, introduces the concept of a reference function, and presents several examples of such functions. Section 3 states the axioms imposed on solutions and the properties required of reference functions, and proves the existence of solutions satisfying these axioms when the reference functions satisfy these properties. The rest of the section studies when the reference functions introduced in Section 2 do satisfy these properties. Section 4 establishes an existence result for multi-valued solutions.

2. REFORMULATION OF THE INDEPENDENCE AXIOM

2.1. *Preliminaries, Bargaining Problems, Solutions, Nash's and Roth's Solutions*

A two-player *bargaining problem* is a pair (S, d), where S is a compact and convex subset of R^2 representing the utility vectors, measured in von Neumann–Morgenstern scales, attainable by the two players through some joint action, and d is a point of S, strictly dominated by some other point of S; d is designated as the "status quo" or the "disagreement point" and is interpreted as the outcome that would result if the players failed to reach a compromise. Σ is the class of all such pairs.

A *solution on* $\Sigma' \subset \Sigma$ is a function associating to every (S, d) in Σ' a point $f(S, d)$ of S called the *solution outcome* for (S, d). This point represents the agreement reached by the players. (Alternatively, one can think of an impartial arbitrator whose role is to select a point of S and to impose it on the players.)

The *Nash solution* [4] is the only solution on Σ satisfying the following four axioms: Pareto-optimality (PO): it should not be possible to increase one agent's utility from the solution outcome without decreasing the other's; *Symmetry* (Sy): if a bargaining problem is symmetrical with respect to the 45° line, the solution outcome should be on that line; *Invariance with respect to positive affine transformation* (Inv): subjecting a bargaining problem to a positive affine transformation and computing the solution outcome are two operations that commute; *Independence of Irrelevant Alternatives* (IIA): if the solution outcome of a bargaining problem is feasible for a subproblem which has the same disagreement point, then it should also be the solution outcome of the subproblem. The *Nash solution outcome* for (S, d) is the unique point maximizing the product $(x - d_x)(y - d_y)$, where $d = (d_x, d_y)$, over the points of S such that $x \geqslant d$.

Introducing the "point of minimal expectations" $m(S) \equiv (b_x(S), a_y(S))$ where $a(S)$ (resp. $b(S)$) is the Pareto-optimal point of S the most favorable to Player 1 (resp. Player 2), Roth reformulated IIA as: *Independence of Alternatives other than* $m(S)$ (IAm), replacing the hypothesis that the two bargaining problems have the same disagreement point by the hypothesis that they have the same point of minimal expectations, and established the existence of a unique solution on Σ satisfying PO, Sy, Inv, and IAm: it is the Nash solution with $m(S)$ as status quo if $m(S)$ is not Pareto-optimal, and $m(S)$ otherwise.

Roth also showed that there is no solution on Σ satisfying PO, Sy, Inv and *Independence of Alternatives other than* $i(S)$ (IAi), an axiom obtained from IAm by replacing $m(S)$ by $i(S) \equiv (a_x(S), b_y(S))$. The "ideal point" $i(S)$ represents the most optimistic aspirations of both players. It appears in the construction of various solutions discussed by Raiffa [5], Luce and Raiffa [2], Kalai and Smorodinsky [3], and Rosenthal [6].

It is the purpose of this paper to propose a general reformulation of the Independence axiom that would encompass Nash's as well as Roth's two formulations, to state conditions under which, in conjunction with PO, Sy, and Inv, this new axiom would imply existence and uniqueness of a solution and to characterize such a solution; finally, to elucidate the negative result described in the preceding paragraph and to offer a natural extension of the concept of a solution that turns this negative result into a positive one.

2.2. *Reference Functions*

The main idea behind the Independence axiom is that the geometry of a bargaining problem is not relevant in all of its details; only a few essential features matter and it is these features that a *reference function* is meant to capture: a reference function $g: \Sigma \to R^2$ associates to every bargaining problem (S, d) in Σ a point $g(S, d)$ of the payoff space, the Independence axiom being formulated with respect to this point. The role of the reference function is essentially informational. In the evaluation of their respective bargaining strengths, the players summarize what they see as the main features of the bargaining problem by focusing on particular outcomes, like the status quo (as in Nash's formulation); or the outcomes that are the most favorable to each of them (as in Roth's two formulations); or both the most and least favorable to each of them; or by adopting some "measure" of the relative sizes of the subsets of feasible alternatives which look "favorable" to eachof them; or by combining any of these considerations with one another and possibly others. $g(S, d)$ is a "reference point" to which the players find it natural to compare any proposed compromise. We will propose here a variety of examples of reference functions.

Before describing them, we observe that the argument just developed does not require that $g(S, d)$ be a feasible alternative. In fact the ideal point, discussed earlier, is feasible only in the trivial case in which it is the only Pareto-optimal outcome, and the point of minimal expectations may not be feasible either.

2.3. *Examples of Reference Functions*

An important way to motivate the search for new reference functions is to exhibit examples of bargaining problems for which the existing solutions do not seem to provide adequate compromises. We will do that explicitly for the Nash solution only.

Given a list $z^1, \ldots, z^n$ of points of R^2, $H(z^1, \ldots, z^n)$ denotes their convex hull. In the examples below, the status quo is taken to be $z^0 = (0, 0)$.

Let now $z^1 \equiv (1, 0)$, $z^2 \equiv (0, 1)$, $S_1 \equiv H(z^0, z^1, z^2)$. The Nash solution outcome for (S_1, z^0) is $z^3 \equiv (1/2, 1/2)$. It was observed by Luce and Raiffa [2] that (S_2, z^0) with $S_2 \equiv (z^0, z^1, z^3, z^4)$ and $z^4 \equiv (0, 1/2)$ has the same Nash solution outcome, which is somewhat disturbing since S_2 is obtained by cutting off from S_1 a region of points that are clearly favorable to Player 2. Player 2's bargaining position seems to have deteriorated in the process and one would like the solution outcome to reflect this change. However, no such move is prescribed by the Nash solution.

The Roth solution is more satisfactory in this respect since it would select z^3 for (S_1, z^0) and $z^5 \equiv (3/4, 1/4)$ for (S_2, z^0). Each coordinate of $m(S)$ represents the maximal utility achievable by the corresponding player when the other player's utility is maximized in S. Assuming that there is enough cooperation among the players to guarantee Pareto-optimality, it is the minimum that each agent can expect, and taking $m(S)$ as reference point reveals a rather pessimistic outlook on the bargaining process. This is why Roth refers to $m(S)$ as the point of minimal expectations.

It is indeed natural to have the points $a(S)$ and $b(S)$, which are the extremities of the Pareto-optimal boundary of S, play a prominent role in the determination of solutions. Another reference function in the definition of which they appear is $M(\cdot)$ which associates to every (S, d) in Σ, the middle $M(S)$ of the segment connecting $a(S)$ and $b(S)$. It seems that the least amount of cooperation among the players would allow them to randomize among $a(S)$ and $b(S)$ thereby achieving $M(S)$. This point could be called the "point of minimal compromise."

Unfortunately, $m(S)$ and $M(S)$ are not continuous functions of S (small changes in S can lead to radical changes in $m(S)$ and $M(S)$), and the solutions associated (in the manner discussed in the introduction) with these reference functions can be criticized for the lack of continuity they inherit. In addition, both solutions are totally irresponsive to the geometry of S in its South, West, and Southwest regions. This should be seen as a weakness since in actual negotiations, the potential losses of the two parties often seem as relevant as their potential gains. A simple reference point that would take both elements into consideration is the center of the smallest rectangle containing S. We will designate it as $D(S)$, and use $D(\cdot)$ as another example of reference function.

An additional example of a reference function is $G(S)$, the center of gravity of S. This reference function has the advantage of depending on the whole of S, and not on accidental features of the boundary of S.

The reference functions examined so far fall into two classes. On the one hand, the Nash reference function is totally insensitive to the geometry of S, and can be criticized on that ground. On the other hand, the reference functions m, M, D, and G are defined independently of the status quo, which is not entirely satisfactory either. A more appealing procedure would take into account both components of the bargaining problem. A simple way to achieve this is to consider linear combinations of the Nash reference function and of one or the other reference functions in the second class. Denoting by β the weight put on the former (and $1-\beta$ that put on the latter), one could say that a high value of β emphasizes the importance of the status quo at the expense of the new opportunities, and therefore generates more conservative reference functions that low values of β.

In the next section, we state two properties of a reference function g which will guarantee the existence of solutions satisfying the four axioms discussed above, the independence axiom being formulated with respect to g. Not surprisingly, the solutions so obtained are related to the Nash solution in the same way as the Roth solution was.

No claim will be made here as to the greatest merit of a particular one of these solutions, the aim of the paper being to propose a wide class of solutions among which one could subsequently discriminate on the basis of additional properties they do or do not have.

Among such properties, it is worth mentioning (a) continuity, how solutions respond to small changes in bargaining problems, (b) "monotonicity," how solutions are affected by finite changes in bargaining problems (for a description and formalization of properties in that spirit, see Thomson and Myerson [7]), (c) generalizability to n-players (all the solutions obtained here pass this test), (d) simplicity and ease of computation (here the solutions involve the determination of the reference point, an easy operation in all cases except perhaps G, and the maximization of a quadratic function).

3. Axioms, Existence Result, Examples

First, we introduce some definitions: Σ_1 is the class of convex and compact subsets of R^2. Σ is the class of bargaining problems: $\Sigma \equiv \{(S, d) \in \Sigma_1 \times R^2 \mid d \in S, \exists z \in S \text{ with } z > d\}$.[1] $\Delta \equiv \{(x, y) \in R^2 \mid x = y\}$ is the 45° line. $T: R^2 \to R^2$ is the *symmetry* (with respect to Δ) operator: given $z = (x, y) \in R^2$, $T(z) = (y, x)$. Given $S \in \Sigma_1$, $T(S) \equiv \{z \in R^2 \mid \exists z' \in S \text{ with } z = T(z')\}$. $A: R^2 \to R^2$ is a *positive affine transformation* if there exists $a_1, a_2 \in R_{++}$, $b_1, b_2 \in R$ such that, given $z = (x, y) \in R^2$, $A(z) = (a_1 x + b_1, a_2 y + b_2)$. $\mathscr{A}$ is the class of these transformations. Given $A \in \mathscr{A}$ and $S \in \Sigma_1$, $A(S) \equiv \{z \in R^2 \mid \exists z' \in S \text{ with } z = A(z')\}$.

Given some domain $\Sigma' \subset \Sigma$, a *solution f on Σ'* associates to every $(S, d) \in \Sigma'$ a point $f(S, d)$, called the solution outcome for (S, d), and interpreted as the compromise reached by the two players when they face the bargaining problem (S, d). The axiomatic approach initiated by Nash [4] consists in imposing axioms on f and in checking whether they are compatible. We follow this approach here, using three of Nash's axioms (PO, Sy, Inv) as well as a generalization of his fourth one (IIA) based on the concept of a reference function discussed in the preceding section.

Let (S, d) and (S', d') be arbitrary elements of Σ'.

Pareto optimality (PO): $f(S, d) \in P(S) \equiv \{z \in S \mid \forall z' \in R^2, z' \geqslant z \Rightarrow z' \notin S\}$.

Symmetry (Sy): If $T(S) = S$ and $g(S, d) \in \Delta$ then $f(S, d) \in \Delta$.

Invariance with respect to positive affine transformations (Inv): For all $A \in \mathscr{A}$, $f(A(S), A(d)) = A(f(S, d))$.

Let $g: \Sigma \to R^2$ be a reference function.

Independence of alternatives other than $g(S, d)$ (IAg): If $S \subset S'$, $g(S, d) = g(S', d')$, and $f(S', d') \in S$, then $f(S, d) = f(S', d')$.

We will be looking for solutions satisfying these four axioms on the domain Σ_g defined by $\Sigma_g \equiv \{(S, d) \in \Sigma \mid \exists z \in S, z > g(S, d)\}$.

We require that g satisfy the following two properties.

Property 1. Invariance with respect to positive affine transformations: For all $A \in \mathscr{A}$, $g(A(S), A(d)) = A(g(S, d))$.

Let $L(S)$ be the unique line of slope -1 supporting S at a Pareto-optimal point, and let $c(S) = L(S) \cap \Delta$.

Property 2. Invariance with respect to symmetrization of almost symmetric bargaining problems: If $g(S, d) \in \Delta$ and $c(S) \in S$, then there exists $(S', d') \in \Sigma_g$ with $S' \supset S$, $T(S') = S$, $c(S') = c(S)$ and $g(S', d') = g(S, d)$.

---

[1] Given vectors x and y in R^2, $x \geqq y$ means that $x_i \geqslant y_i$ for $i = 1, 2$. $x \geqslant y$ that $x \geqq y$ and $x \neq y$, and $x > y$ that $x_i > y_i$ for $i = 1, 2$.

Property 1 is easy to interpret, and should be seen in conjunction with Inv. Property 2 is a technical condition whose role will be clear in the proof of Proposition 1, since it allows to use the same method of proof as Nash's. It states that bargaining problems exhibiting enough symmetry can be replaced by symmetric ones with the same essential features. Although alternative formulations are conceivable, it is clear that a condition of this type cannot be dispensed with, as illustrated by the following example:

$n(S, d)$ is defined to be the middle of the segment connecting $m(S)$ and $a(S)$. Such a function could result from a non-symmetric assessment of the bargaining situations by the two players. Then it would be unreasonable to expect that a solution satisfying PO, Sy, Inv, and IAn would enjoy any symmetry property. Indeed, since for a symmetric bargaining problem (S, d) in Σ, $n(S, d)$ belongs to Δ only if $a(S) = b(S)$, which implies that $P(S)$ is a singleton and that $n(S) = \{P(S)\}$, the hypotheses under which the symmetry axiom applies are never satisfied on Σ_n and therefore the axiom is irrelevant. Multiplicities of solutions should therefore be expected, most of which failing to be symmetric.

Another example illustrating the significance of P2 is provided after Lemma 4. (A class of non-symmetric solutions to the bargaining problem is studied by Brito *et al.* [1].)

We are now ready to prove:

PROPOSITION 1. *If* $g: \Sigma \to R^2$ *satisfies Properties* 1 *and* 2, *there exists a unique solution* f *on* Σ_g *satisfying PO, Sy, Inv, and IAg: it is the Nash solution with* $g(S, d)$ *as status quo.*

Proof. It essentially follows Nash's argument and we write it out for completeness. First, note that the solution described in the statement of the Proposition is well defined since by definition of Σ_g, $g(S, d)$ is strictly dominated by a point of S, for all $(S, d) \in \Sigma_g$.

Sufficiency is clear.

To prove necessity, let $(S, d) \in \Sigma_g$ be given, and let z^* be the maximizer in $P(S)$ of the product $(x - g_x(S, d))(y - g_y(S, d))$. Then, let $A \in \mathscr{A}$ be such that $A(z^*)$ and $A(g(S, d))$ belong to Δ. We have $c(A(S)) = A(z^*) \in \Delta$. By Property 2, there exists $(S', d') \in \Sigma_g$ such that $S' \supset A(S)$, $T(S') = S'$, $c(S') = c(A(S))$ and $g(S', d') = g(A(S), A(d))$. By Property 1, $g(A(S), A(d)) = A(g(S, d))$. Therefore $g(S', d') \in \Delta$. Then, by PO and Sy, $f(S', d') = A(z^*)$. By IAg, $f(A(S), A(d)) = A(z^*)$. By Inv, $f(S, d) = z^*$.

Q.E.D.

Given Proposition 1, in order to check the existence of bargaining solutions associated with the various reference functions described in Section 1, it is enough to show that they satisfy P1 and P2. This is the object of the following lemmas. Making sure that P1 holds is in general an easy task. The verification of P2 requires a more complicated constructive argument. The Lemmas as well as Proposition 2 of the next section are given without proof.

To simplify the verification of P1, we first state:

LEMMA 1. *The reference function* $g_1: \Sigma \to R^2$ *defined by* $g(S, d) = d$ *for all* (S, d) *in* Σ *satisfies Property* 1. *If the reference fucntions* g_1, $g_2: \Sigma \to R^2$ *satisfy Property* 1, *so does the reference function* $\beta g_1 + (1-\beta) g_2$, *for all* β *in* $[0, 1]$..

(a) We now turn to the specific examples discussed in Section 2. The first one is that proposed by Roth. $m(S)$ is the "point of minimal expectations." It is defined by $m(S) = (b_x(S), a_y(S))$, where $a(S)$ and $b(S)$ are the Pareto-optimal points of S the most favorable to Players 1 and 2, respectively.

Then we have

LEMMA 2. *The reference function* $g: \Sigma \to R^2$ *defined, for all* (S, d) *in* Σ, *by* $g(S, d) = m(S)$, *satisfies Properties* 1 *and* 2.

Remark 1. For no β in $]0, 1[$ does the function defined, for all (S, d) in Σ by $g(S, d) = \beta d + (1-\beta) m(S)$ satisfy Property 2.

(b) Next, we consider $M(S)$, the "point of minimal compromise," defined by $M(S) = (a_x(S), b_y(S))$, and we have:

LEMMA 3. *For all* β *in* $[0, 1]$, *the reference function* $g: \Sigma \to R^2$ *defined, for all* (S, d) *in* Σ *by* $g(S, d) = \beta d + (1-\beta) M(S)$, *satisfies Properties* 1 *and* 2.

(c) The following example is a generalization of the reference function $D(\cdot)$ discussed in Section 2. Recall the definition of $D(S)$: we first introduce $x_1 = \min\{x \mid (x, y) \in S\}$, $x_2 = \max\{x \mid (x, y) \in S\}$, $y_1 = \min\{y \mid (x, y) \in S\}$ and $y_2 = \max\{y \mid (x, y) \in S\}$. Then $D(S) \equiv ((x_1 + x_2)/2, (y_1 + y_2)/2)$. This is the smallest rectangle with sides parallel to the axes containing S.

To generalize $D(\cdot)$ we introduce the parameter α in $[0, 1]$, and we define $D_\alpha(S) = (x_\alpha(S), y_\alpha(S))$ where $x_\alpha(S) = \alpha x_1 + (1-\alpha)x_2$ and $y_\alpha(S) = \alpha y_1 + (1-\alpha) y_2)$. Note that for $\alpha = 1/2$, we obtain $D(\cdot)$. The coefficient α reveals the importance given by the players to the worst outcomes in comparison to the best outcomes. We have:

LEMMA 4. *For all* β *in* $[0, 1]$, *and for all* α *in* $[0, 1/2]$, *the reference function* $g: \Sigma \to R^2$ *defined, for all* (S, d) *in* Σ, *by* $g(S, d) = \beta d + (1-\beta) D_\alpha(S)$, *satisfies Properties* 1 *and* 2.

Remark 2. Given any α_1, α_2 in $]0, 1[$, with $\alpha_1 \neq \alpha_2$, the reference function D_{α_1,α_2} defined, for all (S, d) in Σ by $D_{\alpha_1,\alpha_2}(S, d) = \{\alpha_1 x_1 + (1-\alpha_1) y_1, \alpha_2 x_2 + (1-\alpha_2) y_2\}$ which results from a non-symmetric assessment of the bargaining situation by the two players, does not satisfy P2, and the result of Proposition 1 does not apply. This is another example justifying the imposition of P2 on reference functions (see the discussion folowing the statement of P2).

Remark 3. For no α in $]1/2, 1[$ does g satisfy P2 if $\beta \neq 1$. (If $\beta = 1$, g becomes identical to the Nash reference function.)

(d) As a last example, we consider $G(S)$, the center of gravity of S. If S is a degenerate bargaining problem, i.e., a segment, then $G(S)$ is the middle of the segment. Otherwise S is a two dimensional area and $G(S)$ is defined by the equations

$$\iint_S (x - G_x)\, dx\, dy = 0; \qquad \iint_S (y - G_y)\, dx\, dy = 0.$$

We have

LEMMA 5. *For all β in $]0, 1[$, the reference function $g\colon \Sigma \to R^2$ defined, for all (S, d) in Σ, by $g(S, d) = \beta d + (1-\beta)\, G(S)$, satisfies Properties* 1 *and* 2.

Remark 4. By imposing a minor domain restriction, the case $\beta = 0$ can also be dealt with. In the proof of Lemma 5, difficulties arise only for non-symmetric triangles S with a side contained in $L(S)$. By eliminating from consideration all bargaining problems (S, d) for which S has this property, (this constitutes a very minor domain restriction), P2 can be proved to hold even if $\beta = 0$.

Remark 5. Another reference function worth considering associates to every (S, d) in Σ the center of gravity of the boundary of S. Unfortunately, this reference function does not satisfy P1. By using the center of gravity of the Pareto-optimal boundary of S instead, both P1 and P2 can be shown to hold. (I owe this observation to V. Crawford.)

Remark 6. We close this section by observing that the symmetry axiom could have been written as

Symmetry (Sy′): If $T(S) = S$ and $d \in \Delta$, then $f(S, d) \in \Delta$.

This formulation has the advantage of being identical to the one appearing in Nash's paper, and it is also used by Roth. However, since it is the importance of $g(S, d)$ in the evaluation of bargaining problems which is emphasized by IAg, it is that point and not d that should probably be distinguished in the symmetry axiom. If g is the Nash reference function, Sy reduces to Nash's symmetry axiom. By replacing P2 by P2′ below, the same characterization as in Proposition 1 would be obtained. In addition, the various examples of reference functions studied in this paper satisfy P2 when and only when they satisfy P2′. The two formulations are therefore equivalent while we find Sy slightly preferable to Sy′ from the viewpoint of interpretation.

P2′. If $g(S, d) \in \Delta$ and $c(S) \in S$ then there exists $(S', d') \in \Sigma$ with $S' \supset S$, $T(S') = S'$, $d' \in \Delta$, $c(S') = c(S)$ and $g(S', d') = g(S, d)$.

4. MULTI-VALUED SOLUTIONS

As proved in [7] (and mentioned in Section 2), the solution proposed by Roth can be uniquely extended from Σ_m to Σ by noting that (S, d) belongs to $\Sigma \backslash \Sigma_m$ if and only if $m(S)$ is the unique Pareto-optimal point of S. Such an extension cannot be performed for the other examples of reference functions g studied here, but this should not be seen as a serious problem since, for each of them, $\Sigma \backslash \Sigma_g$ is a very small class of games. Turning now to the reference function i (defined in Section 2), we observe, however, that in that case Σ_i is empty. Although this reference function satisfies P1 and P2, [it is immediate to verify that P1 holds; P2 can be shown to hold by using $S' = H(S \cup T(S))$], it is impossible to do away with the requirement that (S, d) be in Σ_i for the conclusion of Proposition 2 to apply: as noted in the introduction, Roth proved, by way of an example (S_0, z_0), that PO, Sy, Inv, and IAi are inconsistent on Σ. A close examination of this example shows that what causes problems is the possibility that there exist several points z of S_0 where S_0 has a line of support with a slope equal to the negative of the slope of the line connecting z to $i(S_0)$, and the example is precisely based on this kind of non-uniqueness.

The natural way to extend the solutions of Section 3 is therefore to allow for multi-valued solutions. The object of a multi-valued solution is to narrow down the set of feasible outcomes to just a few among which the final choice will eventually be made. Bargaining or arbitration procedures often involve successive steps, and what follows can be seen as an attempt to formalize the first step. Of course, to be of interest, this elimination should be substantial enough and, in fact, the most worthwhile procedures will be the ones leading to the greatest elimination. Proving the existence of a minimal solution correspondence when the reference function is $i(\cdot)$ is the object of Proposition 2. Before stating it, it is, however, necessary to reformulate the axioms, which were previously written for functions, so that they be applicable to correspondences.

Given a reference function $g: \Sigma \to R^2$, a *multi-valued solution on* $\Sigma' \subset \Sigma$ is a correspondence associating to every (S, d) in Σ' a non-empty subset of S. In axioms Sy and Inv, the transformations T and A, respectively, should be understood to operate on the *set* $f(S, d)$.

Pareto-optimality (PO): $f(S, d) \subset P(S)$.

Symmetry (Sy): If $T(S) = S$ and $g(S, d) \in \Delta$, then $T(f(S, d)) = f(S, d)$ and $\Delta \cap f(S, d) \neq \varnothing$.

Invariance with respect to positive affine transformations (Inv): For all $A \in \mathcal{A}$, $f(A(S), A(d)) = A(f(S, d))$.

Independence of alternatives other than $g(S, d)$: For all z in $f(S', d')$, if $S \subset S'$, $g(S, d) = g(S', d')$, and $z \in S$, then $z \in f(S, d)$.

Note that if f is required to be single-valued, these axioms become equivalent to these stated in Section 2. We then have:

PROPOSITION 2. *For all β in $[0, 1]$, there exists a multi-valued solution on Σ satisfying PO, Sy, Inv and IAg, where $g: \Sigma \to R^2$ is the reference function defined, for all (S, d) in Σ, by $g(S, d) = \beta d + (1 - \beta)\, i(S)$. There also exists a smallest such solution: it associates to every (S, d) in Σ the set of Pareto-optimal points z of S where S has a line of support of slope different from 0 and ∞, and equal to the negative of the slope of the line connecting $g(S, d)$ and z, if $g(S, d)$ is not Pareto-optimal, or the unique point $g(S, d)$ otherwise.*

REFERENCES

1. D. L. BRITO, A. M. BUONCRISTIANI, AND M. D. INTRILIGATOR, A new approach to Nash's bargaining problem, *Econometrica* **45** (1977), 1163–1172.
2. R. D. LUCE AND H. RAIFFA, "Games and Decisions," Wiley, New York, 1957.
3. E. KALAI AND M. SMORODINSKY, Other solutions to Nash's bargaining problem, *Econometrica* **43** (1975), 513–518.
4. J. F. NASH, The bargaining problem, *Econometrica* **18** (1950), 155–162.
5. H. RAIFFA, Arbitration schemes for generalized two person games, *in* "Annals of Mathematic Studies" (Kuhn and Tucker, Eds.), Vol. 28, Princeton Univ. Press, Princeton, N. Y., 1953.
6. R. ROSENTHAL, An arbitration model for normal-form games, *Math. Operations Res.* **1** (1976), 82–88.
7. A. E. ROTH, Independence of irrelevant alternatives and solutions to Nash's bargaining problem, *J. Econ. Theory* **16** (1977), 247–251.
8. W. THOMSON AND R. B. MYERSON, Monotonicity and independence axioms, *Internat. J. Game Theory* **9** (1980), 37–49.

JOURNAL OF ECONOMIC THEORY 25, 442–449 (1981)

A Nonsubstitution Theorem with Many Primary Factors*

R. MANNING

Department of Economics, University of Canterbury,
Christchurch, 1, New Zealand

Received May 16, 1979; revised October 2, 1980

INTRODUCTION

Johansen [1] presented a neat proof of the celebrated nonsubstitution theorem of Samuelson [7], and extended the concept to the case of joint production, about which he proved that nonsubstitution holds over cones contained in the output space. Only one primary factor was admitted. The present note shows how the earlier results may be further extended, to the case of many primary factors. The theorem proved here is sufficiently general that well-known results on the flatness of production possibility frontiers, as well as nonsubstitution theorems when there is a single primary factor, emerge as special cases. This note provides a link between previously separated branches of the literature.

Nonsubstitution theorems for one primary factor are characterized by, first, the (perhaps only local) independence of the most efficient techniques from final outputs, and secondly, the independence of these techniques from the supply of the primary factor. Johansen's concept of nonsubstitution over a cone exhibits both of these characteristics. An extended definition of nonsubstitution over a cone is developed here. This extension preserves the first characteristic intact, and modifies the second to require that the most efficient techniques are independent of proportional changes in all primary factor supplies (which are the only sort of changes which can occur in primary factor supplies when there is a single primary factor).

If nonsubstitution holds over a cone, then the intersection of this cone with the production possibility frontier is a set which is flat, in the sense that line intervals can be contained in it. Recently, Kemp et al. [4], following up earlier work of Khang and Uekawa [2], and Kemp et al. [3], have given

---

* I am grateful to Murray Kemp, and a referee, for comments on earlier drafts of this note.

0022-0531/81/060442–08$02.00/0

Copyright © 1981 by Academic Press, Inc.

results on the dimension of flat segments of the production possibility frontier. The present analysis complements [4], by explaining the contribution that nonsubstitution makes to the flatness of the production possibility frontier. The technological assumptions made in [4] result in the full employment of all primary factors. Since no such conditions are here imposed, indeed the cones on which nonsubstitution holds are defined with respect to the factors which are fully employed, a generalization of [4] is obtained.

ANALYSIS

There are two kinds of commodity: Produced goods, in number n, and primary factors, in number m. Primary factor j is in fixed total supply L_j, $j = 1,\dots, m$. Produced goods and primary factors are used in production processes by producers, of whom there are r. Producer i has a production set $Y_i \subset R^{n+m}$, Y_i a convex cone with vertex 0, $i = 1,\dots, r$. These assumptions ensure that the production possibility set of the economy is convex. Let $T_i \subset Y_i$ be such that, for all $(\alpha_i, \beta_i) \in Y_i$, there exists a unique $s_i \geqslant 0$ and $(a_i, b_i) \in T_i$ such that $s_i(a_i, b_i) = (\alpha_i, \beta_i)$. T_i is a normalized form of the production set. A production process for producer i is a vector $(a_i, b_i) \in T_i$, for which $a_{ij} > 0$ if and only if produced commodity j is a net output of this process, $a_{ij} < 0$ if and only if produced commodity j is a net input to this process, $a_{ij} = 0$ if and only if produced commodity j is neither a net input nor a net output to this process, and for which $b_{ij} \geqslant 0$, for all $j = 1,\dots, m$, with $>$ for at least one j, so that some primary factors are used as inputs. Let $s_i \geqslant 0$ be the scale of the production process (a_i, b_i). Clearly, $s_i(a_i, b_i) \in Y_i$.

From these definitions,

$$y_j \leqslant \sum_{i=1}^{r} a_{ij} s_i, \qquad j = 1,\dots, n, \tag{1}$$

$$L_j \geqslant \sum_{i=1}^{r} b_{ij} s_i, \qquad j = 1,\dots, m, \tag{2}$$

where y_j is the net output of commodity j for the economy, $j = 1,\dots, n$. In view of these assumptions, in particular the convexity of the individual producers' production sets, it follows that the production possibility set P of the economy is convex. The proof is straightforward.

It is convenient to define the production possibility set through other concepts which help to establish the theorem of this paper.

A *technique* is a collection of production processes $(a_i, b_i) \in I$, for $\{1, \dots, r\}$, and is denoted by (a, b, I). Define the convex cone with verte

$$F(a, I) = \left\{ y \in R^r : y_j = \sum_{i \in I} a_{ij} s_i, \; s_i \geq 0, \; i \in I \right\}$$

using this technique. Within this cone produced goods can be generated a long as the primary factor constraints (2) are not offended.

Let $E \subset \{1, \dots, m\}$, $E \neq \varnothing$, be a *fullemployment set*. Define the set

$$\begin{aligned} M(a, b, I, E) \equiv \Big\{ & y \in R^n_+ : y_j = \sum_{i \in I} a_{ij} s_i; \\ & L_e = \sum_{i \in I} b_{ie} s_i, \; e \in E; \; L_e > \sum_{i \in I} b_{ie} s_i, \; e \notin E; \\ & s_i \geq 0, \; i \in I, \; s_i = 0 \; i \notin I \Big\} . \end{aligned} \tag{4}$$

$M(a, b, I, E)$ is the set of outputs that can be produced with the technology (a, b, I), and with only those resources listed in E fully employed. The set may be empty: It will be empty if there are no solutions $s_i \geq 0$, $i \in I$, to the primary resource constraints given in the definition. In any event, it is obvious that $M(a, b, I, E) \subseteq F(a, I)$.

Let $K(a, b, I, E)$ be the smallest closed cone with vertex 0 containing $M(a, b, I, E)$. This cone is convex, and obviously is contained in $F(a, I)$. When $m = 1$, $K(a, b, I, E) = F(a, I)$, which establishes the connection between the present approach and that of Johansen, who did not need to be concerned with which factors were fully employed.

Define $u(a, b, I, E)$ to be the feasible set of net outputs on the basis of the technique (a, b, I) and the full employment set E. That is,

$$\begin{aligned} u(a, b, I, E) \equiv \Big\{ & y \in R^n_+ : y_j \leq \sum_{i \in I} a_{ij} s_i; \\ & L_e = \sum_{i \in I} b_{ie} s_i, \; e \in E; \; L_e > \sum_{i \in I} b_{ie} s_i, \; e \notin E; \\ & s_i \geq 0, \; i \in I; \; s_i = 0, \; i \notin I \Big\} . \end{aligned} \tag{5}$$

$M(a, b, I, E)$ is the set of efficient points in $u(a, b, I, E)$.

The production possibility set is the union of all feasible sets of ne outputs when all techniques and full employment [illegible]

$$P \quad [illegible] \tag{6}$$

Definition (7). The economy has *nonsubstitution over a cone K* if and only if there is a technique (a^*, b^*, I^*) and a full employment set E^* such that $K \subset K(a^*, b^*, I^*, E^*)$ and, for all a, b, I, E, $u(a, b, I, E) \cap K \subset u(a^*, B^*, I^*, E^*)$. (a^*, b^*, I^*) is a *dominant technique* over the cone K.

This definition of nonsubstitution collapses to that of Johansen if $m = 1$. It retains both features of earlier definitions. Specifically, over $K \cap M(a^*, b^*, I^*, E^*)$ the dominant technique is independent of outputs, while over $K \cap K(a^*, b^*, I^*, E^*)$ the dominant technique is independent of proportional changes in primary factor availability.

Dominant techniques, and associated full employment sets, may be identified by prices.

Definition (8). Nonnegative prices $p_1^*, \ldots, p_n^*$ for produced goods (not all zero), nonnegative prices $w_1^*, \ldots, w_m^*$ for primary factors (not all zero), outputs $y_1^*, \ldots, y_n^*$ of produced goods, the technique (a^*, b^*, I^*), the scales $s_1^*, \ldots, s_r^*$ of production processes, and the full-employment set E^* constitute a *competitive equilibrium* if and only if

$$\sum_{j=1}^{n} a_{ij}^* p_j^* - \sum_{l=1}^{m} b_{il}^* w_l^* = 0, \qquad i \in I^*. \tag{9}$$

$$\sum_{j=1}^{n} a_{ij} p_j^* - \sum_{l=1}^{m} b_{il} w_l^* \leq 0, \qquad \text{for all } (a_i, b_i) \in T_i,\ i \in I^*. \tag{10}$$

$$\sum_{j=1}^{n} a_{ij} p_j^* - \sum_{l=1}^{m} b_{il} w_l^* < 0, \qquad \text{for all } (a_i, b_i) \in T_i,\ i \notin I^*. \tag{11}$$

$$y_j^* \leq \sum_{i \in I^*} a_{ij}^* s_i^*, \qquad j = 1, \ldots, n. \tag{12}$$

$$L_e \geq \sum_{i \in I^*} b_{ie}^* s_i^*, = \text{only if } e \in E^*, \ e = 1, \ldots, m. \tag{13}$$

$$w_e^* > 0 \qquad \text{only if } e \in E^*, \ e = 1, \ldots, m \tag{14}$$

$$p_j^* > 0 \qquad \text{only if} = \text{holds in (12) for } j, \ j = 1, \ldots, n. \tag{15}$$

In this definition (9), (10) and (11) require that maximum profits cannot exceed zero for any producer, and that any technique other than (a^*, b^*, I^*) will cause losses. In particular, (9), (10) and (11) define the members of I^* to be those producers who can avoid losses while remaining in business. Equation (12) specifies that the members of I^* are able to produce the competitive outputs, while (13) requires that their total use of primary resources does not exceed the available quantities. That primary factors must be fully employed before getting a positive price is required by (14), and (15) implies that a produced good has a zero price if less is wanted than is produced. This could happen with joint products.

Zero prices for produced goods cannot occur in equilibrium if, for each produced commodity, there is at least one consumer who prefers more of that commodity. This is a weak assumption of *non satiation*.

THEOREM (16). *If there is a competitive equilibrium with nonsatiation, then* (a^*, b^*, I^*) *is a dominant technique over the cone* $K(a^*, b^*, I^*, E^*)$: *The economy has nonsubstitution over this cone.*

Proof. Suppose otherwise. If the competitive equilibrium technique is not dominant, then in the cone $K(a^*, b^*, I^*, E^*)$ there are some feasible outputs which are better than some competitive equilibrium outputs. That is, there is a technique (a, b, I) and an output vector $y \in K(a^*, b^*, I^*, E^*)$ such that

$$y \gneq y^*, \tag{17}$$

where $y^* \in K(a^*, b^*, I^*, E^*)$ is a competitive equilibrium output vector (and, for vectors, $\gneq$ means that not all components are equal). For some E, $y \in M(a, b, I, E)$.

In (4), multiply y_j by p_j^*, and L_j by w_j^*, and sum. This gives

$$\sum_{j=1}^{n} y_j p_j^* - \sum_{j=1}^{m} L_j w_j^* \leqslant \sum_{j=1}^{n} \sum_{i \in I} a_{ij} s_i p_j^* - \sum_{j=1}^{m} \sum_{i \in I} b_{ij} s_i w_j^*. \tag{18}$$

The right hand side of (18) may be reordered to give

$$\sum_{i \in I} \left(\sum_{j=1}^{n} a_{ij} p_j^* - \sum_{j=1}^{m} b_{ij} w_j^* \right) s_i. \tag{19}$$

But (9), (10) and (11) immediately imply that (19) does not exceed

$$\sum_{i \in I^*} \left(\sum_{j=1}^{n} a_{ij}^* p_j^* - \sum_{j=1}^{m} b_{ij}^* w_j^* \right) s_i^*. \tag{20}$$

On reordering, it is seen that (20) equals

$$\sum_{j=1}^{n} \sum_{i \in I'} a^*_{ij} s^*_i p^*_j - \sum_{j=1}^{m} \sum_{i \in I'} b^*_{ij} s^*_i w^*_j, \tag{21}$$

which, on using (12), (13), (14) and (15), is seen to equal

$$\sum_{j=1}^{n} y^*_j p^*_j - \sum_{j=1}^{m} L_j w^*_j. \tag{22}$$

Hence

$$\sum_{j=1}^{n} y_j p^*_j \leqslant \sum_{j=1}^{n} y^*_j p^*_j. \tag{23}$$

Nonsatiation implies that $p^*_j > 0$, $j = 1,\ldots,n$. (23) therefore contradicts (17). ∎

The definition of a competitive equilibrium does not imply uniqueness of prices: More than one equilibrium price vector can therefore define the same dominant technique. This will happen if the production possibility frontier is not smooth in all directions from a linear segment of the frontier.

The theorem remains true even without nonsatiation, provided that produced commodities have positive prices. If a produced commodity has a zero price it is not certain that (23) and (17) are in conflict, so the theorem cannot be proved. However, little interest attaches to this case, for there is no point in efficiently producing a commodity nobody values.

CONCLUDING COMMENTS

The definitions and result given above collapse to those of Johansen [1] if there is only one primary factor. To illustrate the concepts developed here, and to show that other well known results are special cases of the present theorem, a case with two primary factors is analyzed.

In a well-known paper [6] Melvin showed that the production possibility frontier is a ruled surface if there are three produced goods, no joint products, and two primary factors, with a constant returns to scale, neoclassical production function for each of the industries. These assumptions are a special case of those made earlier, with $n = r = 3$ and $m = 2$.

There are no joint products or intermediate goods, so that $a_{ij} = 1$ or 0 as $i = j$ or $i \neq j$, $i, j = 1, 2, 3$. Then, if all producers operate, so that $I = \{1, 2, 3\}$, $F(a, I) = R^3_+$.

If both primary factors are fully employed $E = \{1, 2\}$, and if a technique

$(a^*, b^*, \{1, 2, 3\})$ is used, then $M(a^*, b^*, \{1, 2, 3\}, \{1, 2\})$ is that part of the intersection of two primary resource constraints which lies in R^3_+. Figure 1 shows R^3_+. The line segment AB is in R^3_+, and is formed by the intersection of the planes defined by the primary resource constraints

$$b^*_{1e} y_1 + b^*_{2e} y_2 + b^*_{3e} y_3 = L_e, \qquad e = 1, 2, \tag{24}$$

on the assumption that the primary factor intensities b^*_{i1}/b^*_{i2}, $i = 1, 2, 3$, are all different. (The primary resource constraints are not shown.)

The cone $K(a^*, b^*, \{1, 2, 3\}, \{1, 2\})$ is formed by the angle AOB.

The five-sided block $ABCDOE$ is the feasible set of net outputs on the basis of the technique $(a^*, b^*, \{1, 2, 3\})$ and the full employment set $\{1, 2\}$. The production possibility set is the union of all such feasible sets of net outputs. In Fig. 1 this is the set defined by $OFGH$, and the line segment AB lies in the ruled production possibility frontier. There is a competitive equilibrium price vector which will cause production to be on AB. For any other technique and full employment set, the feasible set of net outputs $u(a, b, I, E) \cap K(a^*, b^*, \{1, 2, 3\}, \{1, 2\})$ is contained in $u(a^*, b^*, \{1, 2, 3\}, \{1, 2\})$, which illustrates the nonsubstitution theorem.

The method of proof for the present theorem is that used in a slightly different problem by Manning [5, Theorem 1]. The theorem itself is a generalization of that of Johansen [1, Theorem 3.11]. As far as possible, definitions and concepts have been used here in a way that permits comparability with his work. The weakness of the present result, and that of Johansen's also, is that it says that nonsubstitution holds in regions of output space, without specifying the size of these regions: They may be cones without interiors relative to R^n, for instance. If they have interiors relative to

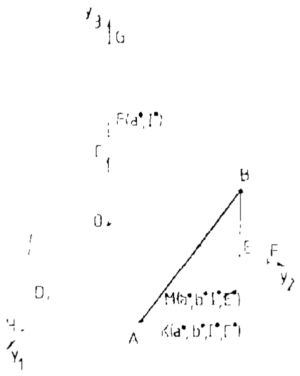


FIG. 1. Illustration of generalized nonsubstitution theorem when $n = r = 3$, $m = 2$ and there are no joint products.

R^n then the result is more forceful. Johansen's remarks on the single primary factor case are pertinent in this regard [1, p. 393–394]. For exact results on the dimensions of flat segments of the production possibility frontier over which nonsubstitution occurs, see Kemp *et al.* [4].

REFERENCES

1. L. JOHANSEN, Simple and general nonsubstitution theorems for input-output models, *J. Econ. Theory* **5** (1972), 383–394.
2. C. KHANG AND Y. UEKAWA, The production possibility set in a model allowing interindustry flows: The necessary and sufficient conditions for its strict convexity, *J. Internat. Econ.* **3** (1973), 283–290.
3. M. C. KEMP, C. KHANG, AND Y. UEKAWA, On the flatness of the transformation surface, *J. Internat. Econ.* **8** (1978), 537–542.
4. M. C. KEMP, R. MANNING, K. NISHIMURA, AND M. TAWADA, On the shape of the single country and world commodity substitution and factor substitution surfaces under conditions of joint production, *J. Internat. Econ.* **10** (1980), 395–404.
5. R. MANNING, Two theorems concerning optimal educational policy in balanced growth, *J. Econ. Theory* **21** (1979), 465–472.
6. J. R. MELVIN, Production and trade with two factors and three goods, *Amer. Econ. Rev.* **58** (1968), 1249–1268.
7. P. A. SAMUELSON, Abstract of a theorem concerning substitutability in open Leontief models, *in* "Activity Analysis of Production and Allocation" (T. C. Koopmans, Ed.), Wiley, New York, 1951.

JOURNAL OF ECONOMIC THEORY 25, 450–452 (1981)

Measuring the Technical Efficiency of Production: A Comment*

RAYMOND J. KOPP

Resources for the Future,
1755 Massachusetts Avenue, N. W., Washington, D. C. 20036

Received August 11, 1979; revised July 23, 1980

In a recent article in this Journal Färe and Lovell [2] (hereafter FL) have proposed a measure of technical efficiency $R(u, x)$ which is applicable under weak production technology assumptions. The objectives of the paper were, "... to specify a set of properties such an efficiency measure should satisfy, to show that the widely used measure proposed by Farrell [3] $F(u, x)$ does not satisfy these properties, and to introduce a new measure that does satisfy these properties."[1]

FL state that $F(u, x)$ satisfies only property E.3. Given production technologies admitting redundant factors, this observation is correct. However, the assumption of strong free disposability of inputs ($x' \geqslant x \Rightarrow \phi(x') > \phi(x)$) is sufficient to ensure that $F(u, x)$ meets all the FL properties E.1–E.4.[2] Since this assumption is contained in the usual set of regularity conditions maintained by most neoclassical–econometric production studies, it does not impose a serious limitation to those studies.

A more fundamental problem concerns the economic interpretation of $R(u, x)$. A well-known property of $F(u, x)$ is its relation to total cost through the formula $1 - F(u, x)$, which is the reduction in total cost associated with the elimination of technical inefficiency. This cost interpretation is independent of both absolute and relative factor prices. On the other hand, $R(u, x)$ cannot be given a meaningful cost interpretation which is factor price invariant. For example, $R(u, x)$ utilizes input vector C (Fig. 1) as its efficient

---

* The patience and helpful comments of C. A. Knox Lovell and the comments of Todd Sandler and V. Kerry Smith are gratefully acknowledged. The author also thanks Rolf Färe for unselfishly providing unpublished manuscripts.

[1] Färe and Lovell [2, p. 150]. The properties designated E.1–E.4 are discussed on p. 157 of Färe and Lovell.

[2] Given $x' \geqslant x \Rightarrow \phi(x') > \phi(x)$ then $\text{Eff}\, L(u) = \text{Iso}\, q(u)$ and consequently $F(u, x)$ satisfies E.1–E.4 (p. 161, Färe and Lovell [2]). A proof of this statement is found in Färe [1, p. 19].

0022-0531/81/060450-03$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

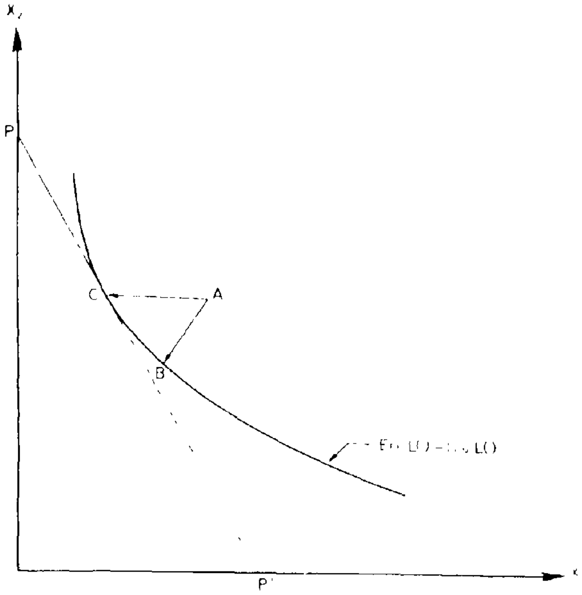


FIGURE 1

reference. Presumably, the ratio of the cost at C to that at A (the inefficient vector) could serve as a measure of technical efficiency that has a cost interpretation. Of course, such a measure would not be equal to $R(u,x)$ but would be more optimistic. More importantly, the cost interpretation would vary with relative factor prices and thus lose the purely physical character of technical efficiency.[3]

The distinction between $F(u,x)$ and $R(u,x)$ is further pronounced if we consider measures of price efficiency. Assuming technical efficiency, the Farrell measure of price efficiency is the ratio of technically efficient total cost to minimum total cost. Constructing the isocost line PP' on Fig. 1, where point C represents the cost minimizing input vector, the price efficiency of vector A is equal to the total cost of vector C divided by the cost of B. If the Farrell price efficiency definition was applied to the FL framework, vector A would be perfectly price efficient, which of course it is not. The ability to successfully dichotomize technical and price efficiency in the Farrell scheme rests upon the radial nature of the efficiency indexes. In contrast, the non-radial $R(u,x)$ does not possess a factor price invariant cost interpretation of technical efficiency, nor can an independent measure of price efficiency be defined.

---

[3] The failure of price invariance can be seen by passing isocost lines through points C and A and then varying the relative prices.

Given strong free disposability, $F(u,x)$ meets the FL criteria and therefore, the choice between $R(u,x)$ and $F(u,x)$ requires additional criteria. Both indexes are capable of ranking production units by a measure of technical efficiency. However, in terms of the economic concept of cost, $F(u,x)$ seems superior. Moreover, at the present time, only the Farrell measurement scheme is capable of incorporating the complementary measure of price efficiency in a consistent fashion.

REFERENCES

1. R. FÄRE, "Laws of Diminishing Returns," Lecture Notes in Economics and Mathematical Systems, p. 19, Springer-Verlag, Berlin, 1979.
2. R. FÄRE AND C. A. K. LOVELL, Measuring the technical efficiency of production, *J. Econ. Theory* **19** (1978), 150–162.
3. M. J. FARRELL, The measurement of productive efficiency, *J. Roy. Statist. Soc. Sect. A* **120** (1957), 253–290.

JOURNAL OF ECONOMIC THEORY 25, 453–454 (1981)

Measuring the Technical Efficiency of Production: Reply

ROLF FÄRE

Department of Economics, Southern Illinois University, Carbondale, Illinois 62901

AND

C. A. KNOX LOVELL

Department of Economics,
University of North Carolina, Chapel Hill, North Carolina 27514

Received June 12, 1980

In [6] we specified a set of properties, E.1–E.4, that a measure of technical efficiency should satisfy. Under axioms ϕ.1–ϕ.5 on the production function ϕ we showed that the Farrell input measure $\mathscr{F}(u, x)$ satisfies only E.3, homogeneity of degree (-1). We then proposed a more general measure, the Russell measure $\mathscr{R}(u, x)$, which satisfies all properties.

In his comment on our paper Kopp [10] suggests replacing our weak disposability of inputs axiom ϕ.3 with a strong free disposability axiom

$$x' \geqslant x \rightarrow \phi(x') > \phi(x). \qquad \phi.3\text{s}$$

If ϕ satisfies this stronger axiom system, plus continuity, then $\mathscr{F}(u, x)$ satisfies E.1–E.4. However, replacing ϕ.3 with ϕ.3s excludes a wide variety of production functions from consideration. For example, ϕ.3s excludes the Leontief production function and a large class of production functions used to model technologies with variable elasticity of substitution (see, e.g., [3, 11, 12]). Moreover, ray homogeneous and ray-homothetic production functions, which permit scale elasticity and hence optimal scale to vary with both the level of output and the factor mix, would be excluded (see [1, 2, 5, 9]). To impose ϕ.3s would also imply that the law of variable proportions would not hold, since the characteristics of congestion necessary for that law to hold would be ruled out (see [4, 8]). Finally, no flexible functional form satisfies ϕ.3s globally.

Kopp also indicates that the Farrell measure is a useful measure of price efficiency when written as $1 - \mathscr{F}(u, x)$. However, it is easy to see that price efficiency measured as $1 - \mathscr{F}(u, x)$ is valid only if ϕ is quasi-concave.

0022-0531/81/060453-02$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

Quasi-concavity, like strong free disposability, excludes from consideration a number of plausible technologies.

Thus Kopp's laudable effort to rescue the Farrell measure, and its price efficiency complement, succeeds only for severely restricted technologies.

We should also mention that Propositions 7 and 8 in [6] have been generalized to read

PROPOSITION. $\mathscr{F}(u, x) \geqslant \mathscr{R}(u, x)$ *for all* $(u, x), x \in L(u), u > 0$.

A proof is available in [7].

REFERENCES

1. W. EICHHORN, "Functional Equations in Economics," Addison–Wesley, Reading, Mass., 1978.
2. R. FÄRE, A note on ray-homothetic and ray-homogeneous production functions, *Swedish J. Econ.* (1975), 366–372.
3. R. FÄRE AND L. JANSSON, On VES and WDI production functions, *Internat. Econ. Rev.* (1975), 745–750.
4. R. FÄRE AND L. JANSSON, Joint inputs and the law of diminishing returns, *Z. Nationalökonomie* (1976), 407–416.
5. R. FÄRE, L. JANSSON AND C. A. K. LOVELL, On ray homothetic production functions, *in* "The Importance of Technology and Permanence of Structure in Industrial Growth," IUI Conference Reports, 2, pp. 228–237, Almquist & Wiksell, Stockholm, 1978.
6. R. FÄRE AND C. A. K. LOVELL, Measuring the technical efficiency of production, *J. Econ. Theory* **19** (1978), 150–162.
7. R. FÄRE, C. A. K. LOVELL, AND K. ZIESCHANG, "Measuring the Technical Efficiency of Multiple Output Production Technologies," presented at a symposium on "Natural Resources and Production" at the University of Karlsruhe, Karlsruhe, West Germany, June 23–July 4, 1980.
8. R. FÄRE AND L. SVENSSON, Congestion of production factors, *Econometrica* (1980), 1745–1753.
9. S. M. GOLDMAN AND R. W. SHEPHARD, On Eichhorn's generalization of homogeneous production functions, *Unternehmensforsch.* (1972), 215–219.
10. R. J. KOPP, Measuring the technical efficiency of production: A comment, *J. Econ Theory* **25** (1981), 450–452.
11. C. A. K. LOVELL, Estimation and prediction with CES and VES production functions, *Internat. Econ. Rev.* (1973), 676–692.
12. N. S. REVANKAR, A class of variable elasticity of substitution production functions, *Econometrica* (1971), 70–71.

JOURNAL OF ECONOMIC THEORY **25**, 455–460 (1981)

Separability vs Strict Separability: A Further Result

ROLF FÄRE AND DANIEL PRIMONT

Department of Economics, Southern Illinois University at Carbondale, Carbondale, Illinois 62901

Received November 5, 1979; revised October 1, 1980

1. INTRODUCTION

In this paper we polish a result by Blackorby, Primont, and Russell (BPR for short, see [1, 2]) on the distinction between "strict separability," as defined by Debreu [4], Stigum [6], and Gorman [5], and "separability," as defined by Bliss [3]. BPR show that "separability" and "strict separability" of a subset of variables in a utility function are equivalent if one assumes that the utility function is locally nonsatiated at every point in its domain with respect to the separable subset. However, this monotonicity condition, while sufficient, is not necessary for the equivalence of the two notions of separability. We offer two alternative, and weaker, conditions, show that they are equivalent, and show that each of them is necessary and sufficient for the equivalence of "separability" and "strict separability."

2. NOTATION AND DEFINITIONS

Our notation here is a subset of that used by BPR. We begin with a continuous, nondecreasing utility function U defined on Ω^n, the nonnegative Euclidean n-orthant. Let $I = \{1, 2,..., n\}$ be the set of integers corresponding to the n components of vectors in Ω^n, and let $\bar{I} = \{I^r, I^c\}$ be a binary partition of I. Corresponding to this partition is a Cartesian decomposition of Ω^n, viz., $\Omega^n = \Omega^{(r)} \times \Omega^{(c)}$. Similarly, a vector X in Ω^n may be written as $X = (X^r, X^c)$.

Each of the two definitions of separability places restrictions on the correspondence β^r, defined by

$$\beta^r(X^r, X^c) := \{\hat{X}^r \in \Omega^{(r)} \mid U(\hat{X}^r, X^c) \geq U(X^r, X^c)\}.$$

0022-0531/81/060455-06$02.00/0

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved

I^r is *strictly separable* from I^c if

$$\beta^r(X^r, X^c) = \beta^r(X^r, \bar{X}^c), \tag{1}$$

for all X^r in $\Omega^{(r)}$ and for all X^c, $\bar{X}^c$ in $\Omega^{(c)}$.

I^r is separable from I^c if

$$\beta^r(X^r, X^c) \subset \beta^r(\bar{X}^r, \bar{X}^c) \qquad \text{or} \qquad \beta^r(\bar{X}^r, \bar{X}^c) \subset \beta^r(X^r, X^c), \tag{2}$$

for all (X^r, X^c), $(\bar{X}^r, \bar{X}^c)$ in Ω^n.

Let

$$N_\delta(\bar{X}^r) := \{X^r \in \Omega^{(r)} \mid (X^r - \bar{X}^r)'(X^r - \bar{X}^r) < \delta^2\}$$

be the usual δ *neighborhood of a point $\bar{X}^r$ in $\Omega^{(r)}$. U is locally nonsatiated with respect to I^r at a point $(\bar{X}^r, \bar{X}^c)$ in Ω^n if for all $\delta > 0$ there exists a point $\hat{X}^r$ in $N_\delta(\bar{X}^r)$ such that $U(\hat{X}^r, \bar{X}^c) > U(\bar{X}^r, \bar{X}^c)$. For brevity we say that U is I^r-locally nonsatiated at $(\bar{X}^r, \bar{X}^c)$.*

3. THE RELATIONSHIP BETWEEN SEPARABILITY AND STRICT SEPARABILITY

In BPR [1] the sufficient condition for the equivalence of separability and strict separability is that U be I^r-locally nonsatiated at *all* $(\bar{X}^r, \bar{X}^c)$ in Ω^n. We replace this by the weaker condition:

> If U is I^r-locally nonsatiated at $(\bar{X}^r, \bar{X}^c)$ then U is I^r-locally nonsatiated at $(\bar{X}^r, X^c)$ for all X^c in $\Omega^{(c)}$. (*)

Thus, we do not require that U be I^r-locally nonsatiated everywhere on the subspace $\Omega^{(r)}$.

If U is I^r locally satiated at some point $(\bar{X}^r, \bar{X}^c)$ in Ω^n then there exists $\delta > 0$ such that for all $\hat{X}^r$ in $N_\delta(\bar{X}^r)$,

$$U(\bar{X}^r, \bar{X}^c) \geqslant U(\hat{X}^r, \bar{X}^c).$$

Consider the alternative condition:

> If U is I^r-locally satiated for some $\delta > 0$ at $(\bar{X}^r, \bar{X}^c)$ then U is I^r-locally satiated for the same $\delta > 0$ at $(\bar{X}^r, X^c)$ for all X^c in $\Omega^{(c)}$. (**)

A simple example of a utility function that satisfies both (*) and (**) is given by:

$$U(x_1, x_2) = (\min(a, x_1)) \cdot x_2.$$

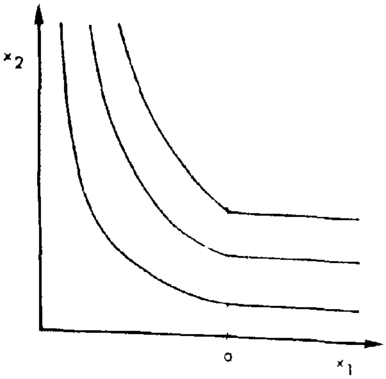


FIG. 1. $U(x_1, x_2) = (\min(a, x_1))\, x_2$.

Some indifference curves of U are sketched in Fig. 1. Both $\{1\}$ and $\{2\}$ are separable subsets of $I = \{1, 2\}$. Suppose $I^r = \{1\}$ and $I^c = \{2\}$. Then, for all $x_2 > 0$, $\beta^r(x_1, x_2)$ is either $[x_1, \infty)$ for $x_1 < a$ or $[a, \infty)$ for $x_1 \geq a$. Thus, $\beta^r(\bar{x}_1, \bar{x}_2) = \beta^r(\bar{x}_1, x_2)$ for all $\bar{x}_1, \bar{x}_2, x_2 > 0$ and I^r is strictly separable from I^c.[1] U is I^r-locally satiated for some $\delta > 0$ (indeed, for all $\delta > 0$) at each point $(x_1, x_2) > 0$ satisfying $x_1 \geq a$ but satisfies both (*) and (**). Thus, I^r local nonsatiation everywhere is not necessary for separability to imply strict separability.

We first prove that:

LEMMA. *(*) and (**) are equivalent conditions.*

Proof. We will show that the negation of either condition implies the negation of the other.

Suppose (*) is false. Then there exists two points $(\bar{X}^r, \bar{X}^c)$ and $(\bar{X}^r, X^c)$, in Ω^n such that U is I^r-locally nonsatiated at $(\bar{X}^r, \bar{X}^c)$ and is I^r-locally satiated at $(\bar{X}^r, X^c)$ for some $\delta > 0$. This implies that (**) is false.

Now suppose (**) is false. There are two relevant ways in which (**) could fail.

(i) U is I^r-locally satiated for some $\delta > 0$ at $(\bar{X}^r, X^c)$ but is I^r locally nonsatiated a $(\bar{X}^r, X^c)$. This implies that (*) is false.

(ii) U is I^r-locally satiated for some $\bar{\delta} > 0$ at $(\bar{X}^r, \bar{X}^c)$, I^r locally satiated for some δ $(0 < \delta < \bar{\delta})$ at $(\bar{X}^r, X^c)$, but is not I^r-locally satiated for $\bar{\delta} > 0$ at $(\bar{X}^r, X^c)$. Let

$$G := \{X^r \mid X^r \in N_{\bar{\delta}}(\bar{X}^r) \text{ and } X^r > \bar{X}^r\}.$$

---

[1] I^r is strictly separable from I^c only for $x_2 > 0$. The example could be easily modified to allow for $x_2 = 0$.

Since U is I^r locally satiated for $\bar{\delta}$ at $(\bar{X}^r, \bar{X}^c)$,

$$U(\bar{X}^r, \bar{X}^c) \geqslant U(\hat{X}^r, \bar{X}^c) \qquad \textit{for all } \hat{X}^r \textit{ in } G.$$

Since U is nondecreasing,

$$U(\bar{X}^r, \bar{X}^c) \leqslant U(\hat{X}^r, \bar{X}^c) \qquad \text{for all } \hat{X}^r \text{ in } G.$$

Thus,

$$U(\bar{X}^r, \bar{X}^c) = U(\hat{X}^r, \bar{X}^c) \qquad \text{for all } \hat{X}^r \text{ in } G.$$

We conclude that:

> U is I^r-locally satiated for some δ' at $(\hat{X}^r, \bar{X}^c)$ for all $\hat{X}^r$ in G (δ' depends on $\hat{X}^r$). (3)

Since U is not I^r-locally satiated for $\bar{\delta}$ at $(\bar{X}^r, X^c)$, and since U is nondecreasing there exists at point $\tilde{X}^r$ in G such that

$$U(\bar{X}^r, X^c) < U(\tilde{X}^r, X^c).$$

Since U is continuous and nondecreasing, this implies that there exists a point $\dot{X}^r$ in G on the line segment connecting $\bar{X}^r$ and $\tilde{X}^r$ such that

$$U \text{ is } I^r \text{ locally nonsatiated at } (\dot{X}^r, X^c) \text{ for some } \dot{X}^r \text{ in } G. \tag{4}$$

Together, (3) and (4) imply that (*) is false.

In both cases (i) and (ii), if (**) is false then (*) is false. ∎

Next, we prove:

THEOREM. *Let U be continuous and nondecreasing. Then I^r is strictly separable from I^c if and only if I^r is separable from I^c and* (*) *holds.*

Proof. *Necessity* (i) strict separability ⇒ separability: obvious from the definitions in (1) and (2).

(ii) strict separability ⇒ (*): Suppose (*) does not hold. Then there exists two points, (X^r, X^c) and $(X^r, \bar{X}^c)$, which differ only in the components of I^c, and a $\delta > 0$ such that

$$U(\hat{X}^r, X^c) > U(X^r, X^c), \tag{5}$$

and

$$U(\hat{X}^r, \bar{X}^c) \leqslant U(X^r, \bar{X}^c), \tag{6}$$

for some $\hat{X}^r$ in $N_\delta(X^r)$. But (5) implies that $X^r \notin \beta^r(\hat{X}^r, X^c)$, while (6) implies that $X^r \in \beta^r(\hat{X}^r, \bar{X}^r)$, contradicting strict separability of I^r. Thus (*) must hold.

Sufficiency: Begin as in BPR by assuming, without loss of generality, that

$$\beta^r(X^r, X^c) \subseteq \beta^r(X^r, \bar{X}^c), \tag{7}$$

by separability of I^r. We need to show that

$$\beta^r(X^r, \bar{X}^c) \subseteq \beta^r(X^r, X^c). \tag{8}$$

Suppose not. Then there exists a point, $\hat{X}^r$, such that

$$\hat{X}^r \in \beta^r(X^r, \bar{X}^c) \tag{9}$$

and

$$\hat{X} \notin \beta^r(X^r, X^c). \tag{10}$$

Consider two cases:

(i) U is I^r-locally nonsatiated at $(\hat{X}^r, \bar{X}^c)$. For this case one may follow the argument in the proof of Theorem 5 in BPR [1] to establish a contradiction.

(ii) U is I^r-locally satiated at $(\hat{X}^r, \bar{X}^c)$. Since case (i) leads to a contradiction U must be I^r-locally satiated at $(\hat{X}^r, \bar{X}^c)$ for all $\hat{X}^r$ in the set difference $\beta^r(X^r, \bar{X}^c) \backslash \beta^r(X^r, X^c)$.

Since $\beta^r(X^r, X^c)$ is closed (continuity of U) and a proper subset of $\beta^r(X^r, \bar{X}^c)$ (from (7), (9) and (10)) there exists a point $\dot{X}^r$, on the boundary of $\beta^r(X^r, X^c)$ which is in $\beta^r(X^r, X^c)$ and in the interior of $\beta^r(X^r, \bar{X}^c)$. For sufficiently small $\delta > 0$ the set,

$$S := \{X^r \mid X^r < \dot{X}^r\} \cap N_\delta(\dot{X}^r),$$

is contained in $\beta^r(X^r, \bar{X}^c) \backslash \beta^r(X^r, X^c)$ since U is nondecreasing. We wish to show that utility is constant on $S \times \{\bar{X}^c\}$.

Consider any path, beginning at $X^r(0)$ in S and ending at $X^r(1) = \dot{X}^r$, such that $X^r(t') > X^r(t)$ if $t' > t$, t', $t \in [0, 1]$. Along such a path $U(X^r(t), \bar{X}^c)$ cannot rise since U is continuous and I^r locally satiated and $U(X^r(t), \bar{X}^c)$ cannot fall since U is nondecreasing. Thus,

$$U(\hat{X}^r, \bar{X}^c) = \bar{k} = \text{constant for all } \hat{X}^r \text{ in } S.$$

Thus, for every $\hat{X}^r$ in S, U is I^r-locally satiated at $(\hat{X}^r, \bar{X}^c)$ for some $\delta > 0$ which depends on $\hat{X}^r$ but, by (**), does not depend on $\bar{X}^c$. Thus, U is I^r-

locally satiated at $(\hat{X}^r, X^c)$ for some $\delta > 0$ for all $\hat{X}^r$ in S. Repeating the argument in the previous paragraph,

$$U(\hat{X}^r, X^c) = k = \text{constant for all } \hat{X}^r \text{ in } S.$$

Moreover since $\dot{X}^r$ is a boundary point of S (from the definition of S) and since U is continuous,

$$U(\dot{X}^r, X^c) = U(\hat{X}^r, X^c) = k \qquad \text{for all } \hat{X}^r \text{ in } S.$$

However, since $\hat{X}^r \in S \subseteq \beta^r(X^r, \bar{X}^c)\backslash\beta^r(X^r, X^c)$ and $\dot{X}^r \in \beta^r(X^r, X^c)$,

$$U(\dot{X}^r, X^c) > U(\hat{X}^r, X^c) \qquad \text{for all } \hat{X}^r \text{ in } S,$$

which is a contradiction.

In both cases, (i) and (ii), a contradiction results. Thus, (8) must be true and therefore I^r is strictly separable. ▮

REFERENCES

1. C. BLACKORBY, D. PRIMONT, AND R. R. RUSSELL, Separability vs functional structure: A characterization of their differences, *J. Econ. Theory* **15** (1977), 135-144.
2. C. BLACKORBY, D. PRIMONT, AND R. R. RUSSELL, "Duality, Separability, and Functional Structure: Theory and Economic Applications," Elsevier/North Holland, New York, 1978.
3. C. J. BLISS, "Capital Theory and the Distribution of Income," North Holland, Amsterdam, 1975.
4. G. DEBREU, Topological methods in cardinal utility theory, *in* "Mathematical Methods in the Social Sciences" (K. Arrow, S. Karlin, and P. Suppes, Eds.), Stanford Univ. Press, Palo Alto, 1959.
5. W. M. GORMAN, The structure of utility functions, *Rev. Econ. Studies* **35** (1968), 369–390.
6. B. P. STIGUM, On certain problems of aggregation, *Internat. Econ. Rev.* **8** (1967), 349–367.

JOURNAL OF ECONOMIC THEORY 25, 461-465 (1981)

A Note on Risk and the Value of Information

YUK SHEE CHAN*

School of Business Administration,
University of California, Berkeley, California 94720

Received December 31, 1979; revised July 2, 1980

This note tries to relate the value of perfect information to the "riskiness" of the decision related environment; "riskiness" is some characteristics of the probability distribution of the states of nature. Assuming implicitly a linear utility function, Gould [1] has shown that value of information generally will be maximized when all but two states are assigned zero probability. We generalize his result for the case of concave utility functions with a more useful definition of value of information.

Suppose $f_i(X)$ is the gross payoff to the decision maker if he has chosen X and state of nature i occurs. X is a control variable defined on an interval $Q \equiv [a, b]$; in particular, we may have $a = -\infty$, $b = \infty$, or both. Assume there are n states of nature; state $i = 1, 2, \ldots, n$ occurs with a probability π_i. Let $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ denote the probability vector; $\boldsymbol{\pi}$ belongs to the set $E = \{\boldsymbol{\pi} \mid 0 \leqslant \pi_i \leqslant 1 \text{ and } \sum_{i=1}^n \pi_i = 1\}$ of probability vectors. We assume expected utility maximization and take utility to be a function of *net* payoff. Further, we assume:

(i) For all i, f_i is concave and differentiable on (a, b); $U(w)$ is defined for all real w and is striclty increasing and concave in w and differentiable everywhere.

(ii) For each probability vector $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$, there exists an interior maximizer $X^*(\boldsymbol{\pi})$ of the expected utility $\sum_{i=1}^n \pi_i U[f_i(X)]$.[1] In particular, if state i is sure to occur, that is, $\pi_i = 1$, we denote the interior maximizer by X_i^*.

The following conditions are satisfied:

$$\sum_{i=1}^{n} \pi_i \frac{\partial U[f_i(X^*(\boldsymbol{\pi}))]}{\partial X} = 0. \tag{1}$$

---

* I am grateful to Professor Thomas Marschak for his encouragement and [illegible] comments.

[1] This assumption is useful in the proofs of Lemma 3 and the Theorem

0022-0531/81/040461-05$02.00/0

Copyright © 1981 by Academic Press, Inc.

$$\frac{\partial U[f_i(X'_i)]}{\partial X} = 0 \qquad \text{for} \quad i = 1, 2, \ldots, n. \tag{2}$$

(iii) We rule out the case where $X'_i = c$ for all i and c is some constant. In this case, the optimal decision is to choose $X = c$ regardless of the value of $\boldsymbol{\pi}$ and the value of information is clearly zero.

Let V denote the maximum amount a decision maker will be willing to pay for perfect information, measured in the same units as payoff is measured. Then V is well defined:[2]

$$\sum_{i=1}^{n} \pi_i U[f_i(X'_i) - V(\boldsymbol{\pi})] = \sum_{i=1}^{n} \pi_i U[f_i(X^*(\boldsymbol{\pi}))], \tag{3}$$

where U is the utility function. Obviously X^* and V are dependent on $\boldsymbol{\pi}$ and we can pose the problem as choosing $\boldsymbol{\pi}$ so as to maximize $V(\boldsymbol{\pi})$ in (3). Clearly, the definition of Gould will coincide with that in (3) if U is linear in its argument. We shall show that Gould's interesting result holds as well with this more useful definition.

Under assumptions (i), (ii), (iii), the following lemmas can be shown in a straightforward manner and except for Lemma 3, are stated without proof.

LEMMA 1. *$V(\boldsymbol{\pi})$ is differentiable on the set of probability vectors $\boldsymbol{\pi}$ and $V(\boldsymbol{\pi}) \geqslant 0$ for all $\boldsymbol{\pi}$.*

Note that a special case of $\boldsymbol{\pi}$ is that $\pi_i = 1$ and $\pi_j = 0$ for $j \neq i$, clearly in this case there is no uncertainty and $V(\boldsymbol{\pi}) = 0$.

LEMMA 2. *There exists $\boldsymbol{\pi}$ with nonzero probability in at least two states such that $V(\boldsymbol{\pi}) > 0$.*

LEMMA 3. *If $\boldsymbol{\pi}^* = (\pi_1^*, \ldots, \pi_n^*)$ is a maximizer of $V(\boldsymbol{\pi})$, then*

$$V(\boldsymbol{\pi}^*) = f_i(X'_i) - f_i(X^*(\boldsymbol{\pi}^*)) \qquad \textit{for all} \quad i \in S, \tag{4}$$

where $S = \{k \mid \pi_k^* > 0\}$ is the set of states with positive probability.[3]

Proof. If $\boldsymbol{\pi}^*$ exists, then Lemma 2 implies that $V(\boldsymbol{\pi}^*) > 0$. But that means that for $\boldsymbol{\pi}^*$ there are at least two states with nonzero probabilities. We can take one of them and denote it state n. For all probability vector $\boldsymbol{\pi}$ with $\pi_n > 0$, we write

---

[2] We are only concerned here with the "buying price" of perfect information. Our definition is the same as that in Marschak and Radner [2]. They have also proved that the value of information as defined in (3) is unique. An alternative "selling price" definition would be the number, Z, of gross payoff units such that best expected utility with perfect information and no deduction from gross payoff equals the best expected utility without information but with Z added to gross payoff.

[3] *We obtain the same interesting result as in Gould's paper.*

$$\pi_n = 1 - \sum_{k=1}^{n-1} \pi_k,$$

and (3) becomes

$$\begin{aligned}\sum_{k=1}^{n-1} \pi_k U[f_k(X'_k) - V(\boldsymbol{\pi})] &+ \left(1 - \sum_{k=1}^{n-1} \pi_k\right) U[f_n(X'_n) - V(\boldsymbol{\pi})] \\ &= \sum_{k=1}^{n-1} \pi_k U[f_k(X^*(\boldsymbol{\pi}))] + \left(1 - \sum_{k=1}^{n-1} \pi_k\right) U[f_n(X^*(\boldsymbol{\pi}))] \qquad (5)\end{aligned}$$

We want to maximize $V(\boldsymbol{\pi})$ in (5) subject to $\pi_k \geqslant 0$ for all $k = 1, 2, \ldots, n-1$.

Differentiate partially with respect to π_i, $i = 1, 2, \ldots, n-1$,

$$\begin{aligned}&U[f_i(X'_i) - V(\boldsymbol{\pi})] - U[f_n(X'_n) - V(\boldsymbol{\pi})] \\ &\quad - \frac{\partial V(\boldsymbol{\pi})}{\partial \pi_i}\left\{\sum_{k=1}^{n} \pi_k U'[f_k(X'_k) - V(\boldsymbol{\pi})]\right\} \\ &= U[f_i(X^*(\boldsymbol{\pi}))] - U[f_n(X^*(\boldsymbol{\pi}))] \\ &\quad + \left\{\sum_{k=1}^{n} \pi_k \frac{\partial U[f_k(X^*(\boldsymbol{\pi}))]}{\partial X}\right\} \frac{\partial X^*(\boldsymbol{\pi})}{\partial \pi_i}. \qquad (6)\end{aligned}$$

Note that the summation terms are from 1 to n.

From (1) the expression in { } in the right-hand side of (6) is zero. Since $U' > 0$, we have for all $i = 1, 2, \ldots, n-1$,

$$\frac{\partial V(\boldsymbol{\pi})}{\partial \pi_i} = \frac{\begin{pmatrix} U[f_i(X'_i) - V(\boldsymbol{\pi})] - U[f_n(X'_n) - V(\boldsymbol{\pi})] \\ - U[f_i(X^*(\boldsymbol{\pi}))] + U[f_n(X^*(\boldsymbol{\pi}))] \end{pmatrix}}{\sum_{k=1}^{n} \pi_k U'[f_k(X'_k) - V(\boldsymbol{\pi})]}. \qquad (7)$$

Applying the Kuhn–Tucker conditions which characterize a maximizer of V on the set of probability vector $\boldsymbol{\pi}$: if $\pi_i^* > 0$, then

$$\frac{\partial V(\boldsymbol{\pi}^*)}{\partial \pi_i} = 0, \qquad \text{for all} \quad i = 1, 2, \ldots, n-1.$$

So for all $i \in S$ (with $n \in S$, as assumed),

$$\begin{aligned}U[f_i(X'_i) - V(\boldsymbol{\pi}^*)] &- U[f_i(X^*(\boldsymbol{\pi}^*))] \\ &= U[f_n(X'_n) - V(\boldsymbol{\pi}^*)] - U[f_n(X^*(\boldsymbol{\pi}^*))]. \qquad (8)\end{aligned}$$

From (3):

$$\sum_{i\in S} \pi_i^* \{U[f_i(X_i') - V(\boldsymbol{\pi}^*)] - U[f_i(X^*(\boldsymbol{\pi}^*))]\} = 0 \tag{9}$$

with $\pi_i^* > 0$ for all $i \in S$, (8) and (9) imply

$$U[f_i(X_i') - V(\boldsymbol{\pi}^*)] = U[f_i(X^*(\boldsymbol{\pi}^*))] \qquad \text{for all} \quad i \in S. \tag{10}$$

We have assumed that U is strictly increasing and continuous on the real line; hence it has a unique inverse. This implies

$$f_i(X_i') - V(\boldsymbol{\pi}^*) = f_i(X^*(\boldsymbol{\pi}^*)) \qquad \text{for all} \quad i \in S,$$

which completes the proof. Q.E.D.

From the proof of Lemma 3, we note that the condition for maximum value of information is (8). Indeed, this condition makes good economic sense with the following interpretation. Given a probability vector $\boldsymbol{\pi}$ and a price V paid for information, the term $d_i = U[f_i(X_i') - V] - U[f_i(X_i^*(\boldsymbol{\pi}))]$ measures the net change in utility if state i obtains. If $d_i > d_j$ and we shift a little bit of probability from state j to i, since information is more valuable in state i, we may expect the decision maker to be willing to pay more for perfect information as a result of the shift of probabilities. The maximum value of information will obtain at the point where d_i is the same for all $i \in S$; this is condition (8). We now have a theorem similar to the one in Gould's paper.

THEOREM.[4] *If $\boldsymbol{\pi}^*$ is a maximizer of $V(\boldsymbol{\pi})$, and there are more than two nonzero probabilities in $\boldsymbol{\pi}^*$, then there is a probability vector which maximizes the value of information and has exactly two nonzero probabilities.*

Proof. From Lemma 3 and (1), $\boldsymbol{\pi}^*$ and $X^*(\boldsymbol{\pi}^*)$ satisfy the following:

$$V(\boldsymbol{\pi}^*) = f_i(X_i') - f_i(X^*(\boldsymbol{\pi}^*)) \qquad \text{for} \quad i \in S \tag{4}$$

and

$$\sum_{i\in S} \pi_i^* \frac{\partial U[f_i(X^*(\boldsymbol{\pi}^*))]}{\partial X} = 0. \tag{11}$$

Note that $\partial U[f_i(X^*(\boldsymbol{\pi}^*))]/\partial X$ cannot be zero for all $i \in S$ because in that case the optimal decision without information is $X = X^*(\boldsymbol{\pi}^*)$ and $V(\boldsymbol{\pi}^*) = 0$.

---

[4] Gould has also proved his Theorem 2 (p. 73) when the control variable is a vector. We shall prove in this note only the case when the control variable is a scalar. The extension to a vector control variable should lead to Gould's Theorem 2.

But by Lemma 2, there exists a probability vector $\boldsymbol{\pi}$ such that $V(\boldsymbol{\pi}) > 0$, so that $\boldsymbol{\pi}^*$ could not be a maximizer of $V(\boldsymbol{\pi})$.

From (11) there must exist two states, call them $l, m \in S$, where $\partial U[f_l(X^*(\boldsymbol{\pi}^*))]/\partial X$ and $\partial U[f_m(X^*(\boldsymbol{\pi}^*))]/\partial X$ are nonzero and have opposite signs.

Let D_l and D_m denote the above two partial derivatives, respectively. Now define $\pi_l, \pi_m > 0$ by

$$\pi_l + \pi_m = 1, \tag{12a}$$

$$D_l \pi_l + D_m \pi_m = 0. \tag{12b}$$

Let $\hat{\boldsymbol{\pi}}$ denote the probability vector $(0, \ldots, 0, \pi_l, 0, \ldots, 0, \pi_m, 0, \ldots, 0)$. From (12b), $X^*(\boldsymbol{\pi}^*)$ also maximizes $\pi_l U[f_l(X)] + \pi_m U[f_m(X)]$ since condition (1) is satisfied. That is,

$$X^*(\boldsymbol{\pi}^*) = X^*(\hat{\boldsymbol{\pi}}). \tag{13}$$

From Lemma 3, since $l, m \in S$,

$$V(\boldsymbol{\pi}^*) = f_l(X'_l) - f_l(X^*(\boldsymbol{\pi}^*)) = f_m(X'_m) - f_m(X^*(\boldsymbol{\pi}^*)). \tag{14}$$

From (13) and (14),

$$V(\boldsymbol{\pi}^*) = f_l(X'_l) - f_l(X^*(\hat{\boldsymbol{\pi}})) = f_m(X'_m) - f_m(X^*(\hat{\boldsymbol{\pi}})). \tag{15}$$

Let $\hat{V} = f_l(X'_l) - f_l(X^*(\hat{\boldsymbol{\pi}})) = f_m(X'_m) - f_m(X^*(\hat{\boldsymbol{\pi}})) = V(\boldsymbol{\pi}^*)$ then

$$f_l(X^*(\hat{\boldsymbol{\pi}})) = f_l(X'_l) - \hat{V}$$

$$f_m(X^*(\hat{\boldsymbol{\pi}})) = f_m(X'_m) - \hat{V},$$

which implies

$$\pi_l U[f_l(X'_l) - \hat{V}] + \pi_m U[f_m(X'_m) - \hat{V}] = \pi_l U[f_l(X^*(\hat{\boldsymbol{\pi}}))] + \pi_m U[f_m(X^*(\hat{\boldsymbol{\pi}}))].$$

By definition, $\hat{V}$ is the value of information for $\hat{\boldsymbol{\pi}}$ and since $\hat{V} = V(\pi^*)$ the proof is complete. Q.E.D.

REFERENCES

1. J. P. GOULD, Risk, stochastic preference, and the value of information, *J. Econ. Theory* **8** (1974), 64–84.
2. J. MARSCHAK AND R. RADNER, "Economic Theory of Teams," Yale Univ. Press, New Haven, 1972.

JOURNAL OF ECONOMIC THEORY 25, 466-470 (1981)

On the Asymptotic Local Stability of Autonomous Economic Systems with Nonnegativity Conditions

GEORGE A. KONDOR*

Department of Economics, Lakehead University,
Thunder Bay, Ontario P7B 5E1, Canada

Received May 27, 1980; revised September 25, 1980

A system of differential equations with nonnegativity conditions—a system of the type used to model price adjustment processes—is proved to be asymptotically locally stable if a subsystem without the nonnegativity conditions is asymptotically locally stable, provided a simple (mathematical) condition prevails.

The asymptotic local stability of an economic system that can be described with the following system of differential equations:

$$\frac{dp_i}{dt} = \begin{cases} f_i[\mathbf{p}(t)] & \text{but} \\ 0 & \text{if } p_i = 0 \text{ and } f_i < 0, \end{cases} \tag{1}$$

where $i = 1,..., n$ and $\mathbf{p}(t) = [p_1(t),..., p_n(t)]$, is examined in this discussion.

Such systems frequently occur in economic literature, e.g., in describing price adjustment processes (this is why we will refer to $p_1,..., p_n$ as prices). A summary of the literature can be found in [3]. The side-constraints are nonnegativity conditions for prices; prices may drop to zero but not below zero. The functions f_i, $i = 1,..., n$, may be related to excess demand. The system is autonomous because f_i does not contain the time variable t explicitly.

The right-hand side of (1) may be discontinuous even if f_i, $i = 1,..., n$, is continuous. This is why the Cauchy–Peano existence theorem does not guarantee a solution. The existence of solutions of similar systems has been dealt with by Henry [2]. It is assumed that f_i $i = 1,..., n$ is continuous and also that system [1] has a solution.

An equilibrium $\bar{\mathbf{p}}$ *for system [1] means a price structure at which* $d\mathbf{p}(t)/dt = \mathbf{0}$. Without side-constraints the local stability analysis of autonomous systems is normally reduced to the investigation of the characteristic values of their "linearized version."[1] If all the characteristic values of

---

* *Present address:* Ontario Hydro, Toronto, Ontario M5G 1X6, Canada.

[1] See [3, pp. 310–311] and also the Poincare-Liapunov Stability Theorem in [1, p. 142

0022-0531/81/060466-05$02.00/0
Copyright © 1981 by Academic Press, Inc.

the linearized model have negative real parts, the original system is asymptotically locally stable (ALS). One may wonder whether a similar simple rule of thumb applies in the presence of the above mentioned side-constraints. The objective of this paper is to show that (1) is ALS at $\bar{\mathbf{p}}$ if the *reduced system*, which does not contain side-constraints and is defined below, is ALS at $\bar{\mathbf{p}}$ provided that a simple condition holds.

We will denote the set of indices $1, \ldots, n$ by I; $i \in I'$ if $i \in I$ and $\bar{p}_i$, the ith equilibrium price, is equal to zero. $i \in I''$ if $i \in I$ and $i \notin I'$.

We will also need the notations in the following definitions:

DEFINITION 1. The *reduced system* is $\boldsymbol{\pi}(t) = [\pi_1(t), \ldots, \pi_n(t)]$, where $\pi_i(t) \equiv 0$ for $i \in I'$, with the adjustment process

$$\frac{d\pi_i(t)}{dt} = f_i[\boldsymbol{\pi}(t)], \qquad i \in I''. \tag{2}$$

DEFINITION 2. The norm of any vector is $\|\mathbf{x}\| = \max_i |x_i|$.

DEFINITION 3. An equilibrium point is called locally stable[2] (or Liapunov stable) at an equilibrium $\bar{\mathbf{p}}$ if, for any positive number α there is a positive number $\gamma(\alpha)$ such that $\|\mathbf{p}(t) - \bar{\mathbf{p}}\| \leqslant \alpha$ for any $t \geqslant 0$, whenever $\|\mathbf{p}_0 - \bar{\mathbf{p}}\| \leqslant \gamma(\alpha)$, where $\mathbf{p}_0$ stands for the initial value.

DEFINITION 4. An equilibrium is called ALS at an equilibrium $\bar{\mathbf{p}}$ if there is a positive number β such that $\mathbf{p}(t) \to \bar{\mathbf{p}}$ as $t \to \infty$ whenever $\|\mathbf{p}_0 - \bar{\mathbf{p}}\| \leqslant \beta$, and the equilibrium also is locally stable.[3]

For the sake of notations that will be used, we also note that the continuity of the function $f_i[\mathbf{p}]$ at $\bar{\mathbf{p}}$ means that for any positive number ε there is a number $\delta = \delta(\varepsilon)$ such that

$$\|\mathbf{f}[\mathbf{p}(t)] - \mathbf{f}[\bar{\mathbf{p}}]\| \leqslant \varepsilon \qquad \text{if} \quad \|\mathbf{p}(t) - \bar{\mathbf{p}}\| \leqslant \delta. \tag{3}$$

CONDITION. *For any* $i \in I'$ $f_i[\bar{\mathbf{p}}] \neq 0$.

This condition may be interpreted such that the excess demand in equilibrium cannot be zero for a commodity whose equilibrium price is zero. Then the excess demand in equilibrium has to be negative.

---

[2] For nonautonomous systems this is the definition of *uniformly* locally stable equilibrium. In case of autonomous systems, like ours, uniform stability implies stability. See [3, pp. 348–350].

[3] *See Footnote 2. The same applies for asymptotically stable systems. It is also worth noting that for autonomous systems asymptotic (local) stability implies (local) stability.*

THEOREM. *If system* (2) *is ALS at* $\bar{\boldsymbol{\pi}}$, *where* $\bar{\pi}_i = \bar{p}_i$ *for all* $i \in I''$, *then system* (1) *is also ALS at* $\bar{\mathbf{p}}$, *provided the above condition holds.*

From this theorem it follows that if the linearized version of the reduced system at $\bar{\boldsymbol{\pi}}$ has characteristic values with negative real parts only, then system (1) will definitely be ALS.

Proof of the Theorem. Consider the values $f_i[\bar{\mathbf{p}}]$ $i \in I'$ and assume that

$$\min_{i \in I'} |f_i[\bar{\mathbf{p}}]| = M. \tag{4}$$

Let us choose the initial value $\mathbf{p}_0$ such that

$$\|\mathbf{p}_0 - \bar{\mathbf{p}}\| \leqslant \frac{M - \varepsilon}{M} \lambda, \tag{5}$$

where $0 < \varepsilon < M$ and there are no other constraints on ε, $\lambda = \min\{\beta, \gamma(\delta), \gamma(\bar{p}_m)\}$, where $\bar{p}_m = \min_{i \in I''} \bar{p}_i$, and the function γ as well as the parameter β are given with reference to system (2). We show that if the initial value of (1) satisfies (5), then $\mathbf{p}(t) \to \bar{\mathbf{p}}$ as $t \to \infty$ provided system (2) is ALS. According to (5), the neighbourhood in which the initial value $\mathbf{p}_0$ is chosen depends on the selection of ε. A sensible choice of ε, in the sense of maximizing the above-mentioned neighbourhood, will be discussed at the end of the paper.

From the Condition we get that $f_i[\bar{\mathbf{p}}] < 0$ for $i \in I'$. Therefore if in (3) ε is chosen such that $\varepsilon < M$, $f_i[\mathbf{p}(t)]$ $i \in I'$ will remain strictly negative while $\mathbf{p}(t)$ is in the δ-neighbourhood of $\bar{\mathbf{p}}$:

$$f_i[\mathbf{p}(t)] < -M + \varepsilon < 0, \qquad i \in I'. \tag{6}$$

For $i \in I''$, $f_i[\bar{\mathbf{p}}] = 0$. Again, from (3)

$$|f_i[\mathbf{p}(t)]| \leqslant \varepsilon \qquad \text{for} \quad i \in I'', \tag{7}$$

provided that $\|\mathbf{p}(t) - \bar{\mathbf{p}}\| \leqslant \delta$.

It can be shown that $\mathbf{p}(t)$ remains in the δ-neighbourhood of $\bar{\mathbf{p}}$ for any t if $\mathbf{p}_0$ is chosen according to (5). We show this step by step.

Starting from its initial price p_{i0} $i = 1, \ldots, n$ the solution $p_i(t)$ $i \in I'$ will get closer to its equilibrium value $\bar{p}_i = 0$ at least for small t values, and then only those prices may move away from their equilibrium prices for which $i \in I''$. This follows from (6).

From (7) and (1) $|dp_i/dt| \leqslant \varepsilon$ $i \in I''$. In other words, ε is at least as large as the absolute value of the rate of change of $p_i(t)$ $i \in I''$, when we are close enough to the equilibrium.

From (5) it is obvious that $p_i(t)$ has to move away from $\bar{\mathbf{p}}$ at least by $\varepsilon(\lambda/M)$ before $p_i(t)$ $i \in I''$ can move out of the λ-neighbourhood of $\bar{\mathbf{p}}$ $|\lambda - ((M - \varepsilon)/M)\lambda = \varepsilon(\lambda/M)|$. To cover this $\varepsilon(\lambda/M)$ distance for prices $i \in I''$, *at least* $\varepsilon(\lambda/M)/\varepsilon = \lambda/M$ time is required. Within this time, however, all $p_i(t)$ $i \in I'$ reach zero as it is shown below:

From (6), for $t \leqslant \lambda/M$, and before $p_i(t)$, $i \in I'$, vanishes, $p_{i0} - p_i(t) \geqslant (M - \varepsilon)t$, or

$$p_{i0} - (M - \varepsilon)t \geqslant p_i(t). \tag{8}$$

The left-hand side of (8) becomes negative or zero at $t_0 = \lambda/M$:

$$p_{i0} - (M - \varepsilon)\frac{\lambda}{M} = (p_{i0} - \bar{p}_i) - \frac{M - \varepsilon}{M}\lambda \leqslant 0 \tag{9}$$

since $\bar{p}_i = 0$ for $i \in I'$. From (9) it follows that $p_i(t)$ has to vanish within time λ/M.

It has been shown that if the process starts as it is prescribed in (5), all $p_i(t)$ $i \in I'$ will reach zero without the system moving out of the λ neighbourhood of the equilibrium price vector. In this λ-neighbourhood (6) and (7) prevail and, due to local stability and the definition of λ, the system will remain in the δ neighbourhood for any $t \geqslant t_0$ as well. Then it follows that $p_i(t)$ $i \in I'$ cannot be different from zero for $t \geqslant t_0$.

The question then arises whether system (2) is equivalent to system (1) for $t \geqslant t_0$ with initial value $\mathbf{p}(t_0)$. It is so if we can show that the solution of system (2) is nonnegative. Since $\lambda \leqslant \gamma(\bar{p}_m)$, due to local stability, it follows that $\|\mathbf{p}(t) - \bar{\mathbf{p}}\| \leqslant \bar{p}_m$, $t \geqslant t_0$ or from the definition of the norm, $\max_i |p_i(t) - \bar{p}_i| \leqslant \min_{i \in I''} \bar{p}_i$, $t \geqslant t_0$. This means that the solution of system (2) is nonnegative for any $t \geqslant t_0$. The system (2) is then equivalent to system (1). Since $\|\mathbf{p}(t_0) - \bar{\mathbf{p}}\| \leqslant \delta$, system (2) is ALS, then so is system (1). Q.E.D.

It has been proved that if the initial solution is chosen within a neighbourhood of the equilibrium

$$\mu = \frac{M - \varepsilon}{M}\min\{\beta, \gamma(\delta), \gamma(\bar{p}_m)\} \tag{10}$$

as specified in (5), then the system (1) is ALS, provided that system (2) is ALS and the Condition prevails. It has not been shown, neither claimed, however, that the process has to start *in the μ neighbourhood of $\bar{\mathbf{p}}$ in order to* converge to $\bar{\mathbf{p}}$. Therefore, the question of how to select ε to maximize μ in

(10) has only limited significance. Still, it may be worth noting that the following nonlinear programming problem answers this question:

$$\max \mu$$

$$\varepsilon \geqslant 0; \qquad \mu \leqslant \frac{M-\varepsilon}{M} \min\{\beta, \gamma(\bar{p}_m)\}, \qquad \text{and} \qquad \mu \leqslant \frac{M-\varepsilon}{M} \gamma[\delta(\varepsilon)].$$

(Of course $\varepsilon < M$ has to prevail. Since $\varepsilon \geqslant M$ cannot be a solution, the nonnegativity of ε is the only additional requirement to be explicitly considered. It is assumed above that $\delta(\cdot)$ is the largest radius for which (3) prevails and $\gamma(\cdot)$ is the largest radius that guarantees the stability of System (2).)

REFERENCES

1. R. BELLMAN, "Methods of Nonlinear Analysis, Vol. 1," Academic Press, New York/San Francisco, London, 1970.
2. C. HENRY, Differential equations with discontinuous right-hand side for planning procedures, *J. Econ. Theory* **4** (1972), 545–551.
3. A. TAKAYAMA, "Mathematical Economics," Dryden, Hinsdale, Ill., 1974.

JOURNAL OF ECONOMIC THEORY 25, 471 (1981)

Erratum

Volume **24**, Number 1 (1981), in the article, "The Overlapping-Generations Model. II. The Case of Pure Exchange with Money" by Yves Balasko and Karl Shell, pp. 112–142:

Monetary economics is a musterious subject. Cf., e.g., Fig. 7.1 on page 136 of our paper on overlapping generations with money. The set $\mathscr{M}_B(\omega)$ is not visible. We offer Fig. 7.1 once again. This time we have "pressed harder." What you should see is the set of normalized ω-bonafide monetary policies and the intersection of that set with the ray $\{\lambda\mu \mid \lambda \geqslant 0\}$.

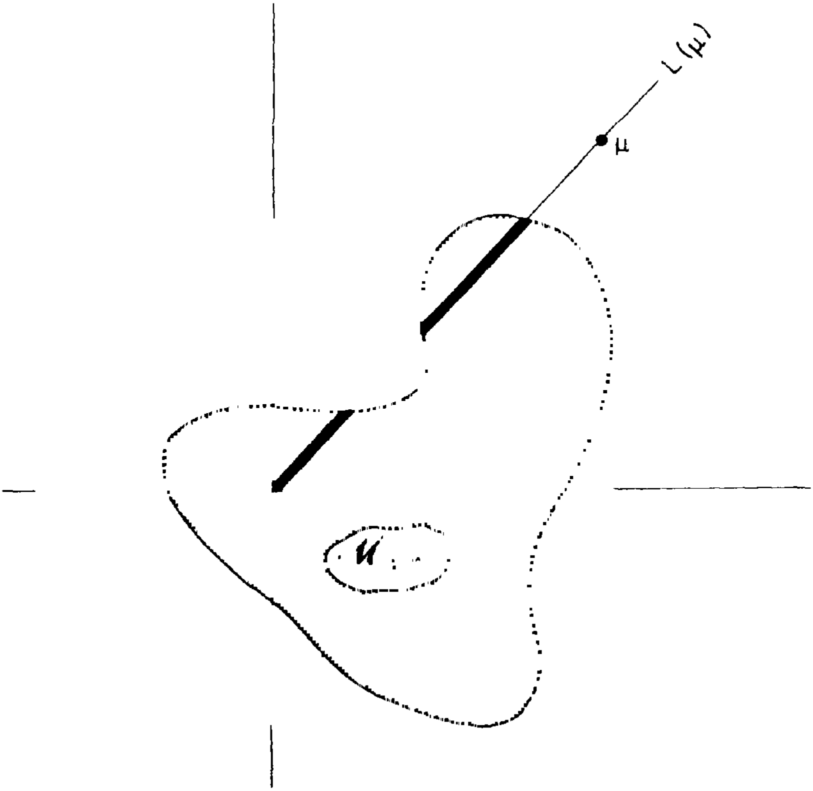


FIG. 7.1

YVES BALASKO
KARL SHELL

0022-0531/81/060471-01$02.00/0
Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

JOURNAL OF ECONOMIC THEORY 25, 472-473 (1981)

Papers to Appear in Forthcoming Issues

The following articles are among those newly accepted for publication in the *Journal of Economic Theory*. They are listed alphabetically by first author.

BETH ALLEN, Approximate Equilibria in Microeconomic Rational Expectations Models. Correspondent: Professor Beth Allen, Department of Economics, University of Pennsylvania, 3718 Locust Walk/CR, Philadelphia, PA 19104

TARADAS BANDYOPADHYAY, Coalitional Manipulation and the Parato Rule. Correspondent: Dr. Taradas Bandyopadhyay, Department of Economics, University of Hull, Hull, HU6 7RX England

GARY CHAMBERLAIN, A Characterization of the Distributions that Imply Mean–Variance Utility Functions. Correspondent: Professor Gary Chamberlain, Department of Economics, University of Wisconsin, Madison, WI 53706.

JOHN B. DONALDSON AND RAJNISH MEHRA, Stochastic Growth with Correlated Production Shocks. Correspondent: Professor John B. Donaldson, Graduate School of Business, Columbia University, New York, NY 10027

JOHN A. FEREJOHN AND RICHARD D. MCKELVEY, Von Neumann Morgenstern Solution Social Choice Functions: An Impossibility Theorem. Correspondent: Professor Richard D. McKelvey, School of Urban and Public Affairs, Carnegie-Mellon University, Pittsburgh, PA 15213.

RICHARD JENSEN, Innovation Adoption and Diffusion when There Are Competing Innovations. Correspondent: Professor Richard Jensen, Department of Economics, Ohio State University, 1775 College Road, Columbus, OH 43210.

DAVID M. KREPS, PAUL MILGROM, JOHN ROBERTS, AND ROBERT WILSON, Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. Correspondent: Professor David M. Kreps, Graduate School of Business, Stanford University, Stanford, CA 94305

DAVID M. KREPS AND ROBERT WILSON, Reputation and Imperfect Information. Correspondent: Professor David M. Kreps, Graduate School of Business, Stanford University, Stanford, CA 94305.

PAUL MILGROM AND JOHN ROBERTS, Predation, Reputation, and Entry Deterrence. Correspondent: Professor John Roberts, Graduate School of Business, Stanford University, Stanford, CA 94305.

JAMES C. MOORE, Measurable Triples and Cardinal Measurement. Correspondent: Professor James C. Moore, Purdue University, 430 Krannert Building, West Lafayette, IN 47907

KEVIN ROBERTS, Social Choice Rules and Real-Valued Representations. Correspondent: Dr. K. Roberts, St. Catherine's College, Oxford, OX1 3UJ England

DAVID SAPPINGTON, Limited Liability Contracts Between Principal and Agent. Correspondent: Professor David Sappington, Department of Economics, University of Michigan, Ann Arbor, MI 48109.

CLAUS WEDDEPOHL, Fixed Price Equilibria in a Multifirm Model. Correspondent: Professor

Copyright © 1981 by Academic Press, Inc.

Claus Weddepohl, Faculty of Actuarial Science and Econometrics, University of Amsterdam, Jodenbreestraat 23, Amsterdam, 1011 NH The Netherlands

LES WITCHARD, Optimal Structures for the Residential Zone of a Factory Town
Correspondent: Professor L. C. Witchard, Department of Economics, Faculty of Economics, Australian National University, P.O. Box 4, Canberra, A.C.T. 2600, Australia

JOURNAL OF ECONOMIC THEORY **25**, 474–475 (1981)

Author Index for Volume 25

A

Avriel, M., 397
Azariadis, Costas, 380

B

Balasko, Yves, 471
Balcer, Yves, 196
Burdett, K., 42

C

Chamberlain, Gary, 152
Champsaur, Paul, 269
Chan, Yuk Shee, 461
Cowell, Frank A., 131

D

Diamond, P. A., 165
Diewert, W. E., 397

E

Eckalbar, John C., 302

F

Färe, Rolf, 453, 455
Fishburn, Peter C., 255

G

Gardner, Roy, 353
Guesnerie, Roger, 67

H

Haque, W., 237
Heiner, Ronald A., 101
Hughes, Joseph P., 287

I

Ichishi, Tatsuro, 283

K

Katz, Eliakim, 318
Kondor, George A., 466
Kopp, Raymond J., 450
Kuga, Kiyoshi, 131

L

Laroque, Guy, 269
Lovell, C. A. Knox, 453

M

MacIntyre, I., 338
Manning, R., 442
Maskin, Eric, 165
Mortensen, Dale T., 42

O

Oddou, Claude, 67

P

Pattanaik, Prasanta K., 338
Polemarchakis, Heraklis M., 323
Primont, Daniel, 455

Copyright © 1981 by Academic Press, Inc.
All rights of reproduction in any form reserved.

R

ADER, TROUT, 219
INGANUM, JENNIFER F., 21
OSENFIELD, DONALD B., 1
OSENTHAL, ROBERT W., 92
OTHSCHILD, MICHAEL, 152

S

HAPIRO, ROY D., 1
HELL, KARL, 471
TARR, ROSS M., 314
UZUMURA, KOTARO, 367

T

THOMSON, WILLIAM, 431

W

WALLACE, NEIL, 421
WOODERS, MYRNA, 144

Z

ZANG, I., 397

JOURNAL OF Economic Theory

EDITOR: Karl Shell

ASSOCIATE EDITORS:

Yves Balasko
William A. Brock
David Cass
Paul Champsaur
Gerard Debreu
Peter C. Fishburn
Richard J. Gilbert
Jean-Michel Grandmont
Sanford Grossman
Roger Guesnerie
Walter Perrin Heller
Ehud Kalai
David M. Kreps
Mordecai Kurz
Jean-Jacques Laffont
Mukul Majumdar
Edmond Malinvaud
Prasanta K. Pattanaik
Andrew Postlewaite
Roy Radner
John Roberts
Stephen A. Ross
Michael Rothschild
Wayne J. Shafer
David A. Starrett
Neil Wallace
Charles A. Wilson

ASSISTANT EDITOR: Susan Witherow Schulze

Volume 25, 1981

ACADEMIC PRESS
A Subsidiary of Harcourt Brace Jovanovich, Publishers
New York London Toronto Sydney San Francisco

Copyright © 1981 by Academic Press, Inc

All Rights Reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc. (21 Congress Street, Salem, Massachusetts 01970), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre 1981 articles are the same as those shown for current articles.

Printed by the St. Catherine Press, Ltd., Bruges, Belgium

Contents of Volume 25

NUMBER 1, AUGUST 1981

DONALD B. ROSENFIELD AND ROY D. SHAPIRO. Optimal Adaptive Price Search 1

JENNIFER F. REINGANUM. Dynamic Games of Innovation . . 21

K. BURDETT AND DALE T. MORTENSEN. Testing for Ability in a Competitive Labor Market 42

ROGER GUESNERIE AND CLAUDE ODDOU. Second Best Taxation as a Game 67

ROBERT W. ROSENTHAL. Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox 92

RONALD A. HEINER. Length and Cycle Equalization 101

NOTES, COMMENTS, AND LETTERS TO THE EDITOR.

FRANK A. COWELL AND KIYOSHI KUGA. Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement 131

MYRNA WOODERS. Equilibria, the Core and Jurisdiction Structures in Economies with a Local Public Good: A Correction 144

GARY CHAMBERLAIN AND MICHAEL ROTHSCHILD. A Note on the Probability of Casting a Decisive Vote 152

PAPERS TO APPEAR IN FORTHCOMING ISSUES 163

NUMBER 2, OCTOBER 1981

P. A. DIAMOND AND ERIC MASKIN. An Equilibrium Analysis of Search and Breach of Contract. II. A Non-Steady State Example 165

YVES BALCER. Equilibrium Distributions of Sales and Advertising Prices over Space 196

TROUT RADER. Utility over Time: The Homothetic Case 219

W. HAQUE. Direct and Indirect Weak Separability 237

PETER C. FISHBURN. Majority Committees 255

PAUL CHAMPSAUR AND GUY LAROQUE. Fair Allocations in Large Economies 269

NOTES, COMMENTS, AND LETTERS TO THE EDITOR

TATSURO ICHIISHI. Super-modularity: Applications to Convex Games and to the Greedy Algorithm for *LP* 283

JOSEPH P. HUGHES. Giffen Inputs and the Theory of Multiple Production 287
JOHN C. ECKALBAR. Stable Quantities in Fixed Price Disequilibrium 302
ROSS M. STARR. Approximation of Points of the Convex Hull of a Sum of Sets by Points of the Sum: An Elementary Approach 314
ELIAKIM KATZ. A Note on a Comparative Statics Theorem for Choice under Risk 318

PAPERS TO APPEAR IN FORTHCOMING ISSUES 320

NUMBER 3, DECEMBER 1981

HERAKLIS M. POLEMARCHAKIS. Constrained Excess Demand Functions 323
I. MACINTYRE AND PRASANTA K. PATTANAIK. Strategic Voting under Minimally Binary Group Decision Functions 338
ROY GARDNER. Wealth and Power in a Collegial Polity 353
KOTARO SUZUMURA. On Pareto-Efficiency and the No-Envy Concept of Equity 367
COSTAS AZARIADIS. Self-Fulfilling Prophecies 380
W. E. DIEWERT, M. AVRIEL, AND I. ZANG. Nine Kinds of Quasiconcavity and Concavity 397
NEIL WALLACE. A Hybrid Fiat–Commodity Monetary System . 421

NOTES, COMMENTS, AND LETTERS TO THE EDITOR

WILLIAM THOMSON. A Class of Solutions to Bargaining Problems 431
R. MANNING. A Nonsubstitution Theorem with Many Primary Factors 442
RAYMOND J. KOPP. Measuring the Technical Efficiency of Production: A Comment 450
ROLF FÄRE AND C. A. KNOX LOVELL. Measuring the Technical Efficiency of Production: Reply 453
ROLF FÄRE AND DANIEL PRIMONT. Separability vs Strict Separability: A Further Result 455
YUK-SHEE CHAN. A Note on Risk and the Value of Information 461
GEORGE A. KONDOR. On the Asymptotic Local Stability of Autonomous Economic Systems with Nonnegativity Conditions 466

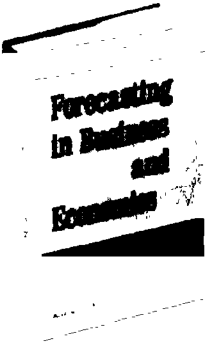
ERRATUM

Volume 24, Number 1, 1981: YVES BALASKO AND KARL SHELL, "The Overlapping Generations Model. II. The Case of Pure Exchange with Money," pp. 112–142 471

PAPERS TO APPEAR IN FORTHCOMING ISSUES 472

AUTHOR INDEX FOR VOLUME 25 474

FORECASTING IN BUSINESS AND ECONOMICS



By C. W. J. GRANGER

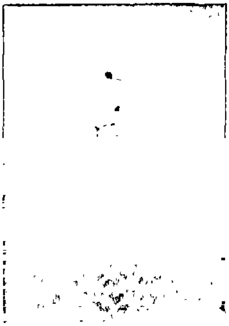
University of California, San Diego

Written by the recognized authority in forecasting theory and application, *Forecasting in Business and Economics* gives the senior, MBA, and beginning graduate student a [illegible] strategy for tackling forecasting problems.

Unlike any other book in the field, this new text [illegible] the importance of selecting a relevant information set and methods of evaluating the forecast once it has been prepared. Granger's emphasis on [illegible] assessment gives the future economist and manager a sound basis for interpreting forecasts and making decisions about the future.

Modern techniques of statistical and econometric [illegible] and technological forecasting are presented along with [illegible] a wide range of topics, including Box–Jenkins methods, [illegible] survey data, world models, and the use of [illegible]

1980, 226 pages



FUNDAMENTALS OF BUSINESS STATISTICS

JEROME D. BRAVERMAN [illegible]

A business statistics [illegible] student [illegible] Student [illegible] [illegible]

A [illegible] book [illegible] with [illegible] background [illegible] succeed in the course. Only [illegible] knowledge of high school math is required.

A *Student* [illegible] available.

[illegible] pages

ACADEMIC PRESS, INC.

A [illegible] Publishers
NEW YORK LONDON TORONTO SYDNEY SAN FRANCISCO

[illegible] ACADEMIC PRESS [illegible] New York, NY [illegible]

AP 7964

How To Comply With The New Copyright Law

Libraries everywhere have found the easy way to fill photocopy requests legally and instantly, without the need to seek permissions, from this and over 3000 other key publications in business, science, humanities, and social science

Participation in the Copyright Clearance Center (CCC) assures you of legal photocopying at the moment of need You can

Fill requests for multiple copies, interlibrary loan (beyond the CONTU guidelines), and reserve desk without fear of copyright infringement

Supply copies simply and easily from registered publications The CCC's flexible reporting system accepts photocopying reports and returns an itemized invoice *You need not keep any records*, our computer will do it for you

The Copyright Clearance Center is your one-stop place for on-the-spot clearance to photocopy for internal use You will never have to decline a photocopy request or wonder about compliance with the law for any publication registered with the CCC

For more information, just contact

Copyright Clearance Center
21 Congress Street
Salem, Massachusetts 01970
(617) 744-3350

a not-for-profit corporation

NAME TITLE

ORGANIZATION

ADDRESS

CITY STATE ZIP

COUNTRY TELEPHONE